Alma Mater Studiorum – Università di Bologna
In collaborazione con LAST-JD consortium:
Università degli studi di Torino
Universitat Autonoma de Barcelona
Mykolas Romeris University
Tilburg University
e in cotutela con
THE Luxembourg University

DOTTORATO DI RICERCA IN

Erasmus Mundus Joint International Doctoral Degree in
Law, Science and Technology

Ciclo XXXI – A.A. 2015/2016

Settore Concorsuale: IUS20
Settore Scientifico Disciplinare: 12H3

LEGAL DESIGN FOR THE GENERAL DATA PROTECTION REGULATION A METHODOLOGY FOR THE VISUALIZATION AND COMMUNICATION OF LEGAL CONCEPTS

Presentata da:
Arianna Rossi

Coordinatore Dottorato	Supervisore
Prof. Giovanni Sartor	Prof.ssa Monica Palmirani
	Prof. Leon van der Torre

Esame finale anno 2019

UNIVERSITÉ DU LUXEMBOURG

PhD-FSTC-2019-19
The Faculty of Sciences, Technology and Communication

DISSERTATION

Presented on 29/03/2019 in Bologna
to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN INFORMATIQUE

by

Arianna Rossi
Born on 24 September 1989 in Arezzo (Italy)

Legal Design for the General Data Protection Regulation
A Methodology for the Visualization and Communication of Legal Concepts

Alma Mater Studiorum – Università di Bologna
in partnership with LAST-JD Consortium
Università degli studi di Torino
Universitat Autonoma de Barcelona
Mykolas Romeris University
Tilburg University
and in cotutorship with
THE University of Luxembourg

PhD Programme in

Erasmus Mundus Joint International Doctoral Degree in
Law, Science and Technology

Cycle XXXI

Settore Concorsuale di afferenza: 12H3
Settore Scientifico disciplinare: IUS20

Legal Design for the General Data Protection Regulation
A Methodology for the Visualization and Communication of Legal Concepts

Submitted by: Arianna Rossi

Supervisors
Prof. Monica Palmirani
Prof. Leon van der Torre

The PhD Programme Coordinator
Prof. Giovanni Sartor

Year 2019

To my parents
Zinaida and Roberto

Acknowledgements

Firstly, I would like to thank all the people and institutions without whom this research would have not been possible:

Stanford University, Legal Design Lab: Margaret Hagan;

Accademia di Belle Arti di Bologna: Prof. Danilo Danisi and (in alphabetical order) Sara Cominelli, Grazia Coppolecchia, Nicoletta Guastalla, Francisca Lessio, Valentina Marafioti, Emy Michelin, Alyssa Scroffa, Vladimir Senyuk, Francesca Velleca, Nikolas Vigorito and Federica Zanasi;

University of Bologna, CIRSFID: (in alphabetical order) Giorgia Bincoletto, Dino Gerardi, Dina Ferrari, Michele Martoni, Maria Livia Rizzo, Michela Rossi and Salvatore Sapienza;

University of Luxembourg: Professor Leon van der Torre.

The research has also been supported by the Marie Sklodowska-Curie Mining and Reasoning with Legal Texts (MIREL) project and I would like to gratefully acknowledge Livio Robaldo, as a colleague and friend. I would also like to thank Chiara Cavazzuti (Accademia di Belle Arti di Firenze) and Claudia Busetto (Commonground Srl) for their valuable contributions. A special thanks goes to Miroslav Kurdov (SketchLex) for his precious help on the final refinement of the icon set and the choice of the CC license. I gratefully acknowledge the EACEA for financing the unique PhD program LAST-JD and the SIIG (Società Italiana Informatica Giuridica). Finally, all

the participants in the "Law & Design for Privacy" workshops and in the user studies are gratefully acknowledged.

I am very grateful to my supervisor, Professor Monica Palmirani: thank you for your great encouragement, for everything that you have taught me and for having introduced me to the academic life, which I love. Without you, this research, and my personal and professional development would have not been possible. I would also like to thank my supervisor Professor Leon van der Torre: thank you for having warmly welcomed me at the University of Luxembourg, for the opportunities that you have offered me and for your continuous support. I am grateful to all the institutions, professors and colleagues that have welcomed my PhD pilgrimage: University of Bologna, University of Turin, Universitàt Autonoma de Barcelona, Stanford University and University of Luxembourg.

I would also like to express all my gratitude and my affection to the colleagues, co-authors, fellow legal designers and now friends that have incredibly contributed to this voyage of discovery: Colette Brunschwig, Rossana Ducato, Helena Haapio, Margaret Hagan, Stefania Passera. I can't express in words all that I have learned from you on a professional and personal level. I wish we will always collaborate and brainstorm together.

I am also very grateful to all my family, without whom I would have never made it: to my parents, to my brother, to my aunts and uncles. Thanks for your continuous support in the motherland and beyond. I also think with affection about all the friends that I have met in these years and those that have always been there, even before.

Lastly, I would like to thank with my deepest love I & I lion.

Contents

List of Figures

List of Tables

Abstract

Privacy policies are known to be impenetrable, lengthy, tedious texts that are hardly read and poorly understood. Therefore, the General Data Protection Regulation (GDPR) introduces provisions to enhance the transparency of such documents and suggests icons as visual elements to provide "in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing." The present dissertation discusses how design, and in particular legal design, can support the concrete implementation of the GDPR's transparency obligation. Notwithstanding the many benefits that visual communication demonstrably provides, graphical elements do not improve comprehension *per se*. Research on graphical symbols for legal concepts is still scarce, while both the creation and consequent evaluation of icons depicting abstract or unfamiliar concepts represent a challenge. Moreover, precision of representation can support the individuals' sense-making of the meaning of graphical symbols, but at the expense of simplicity and usability. Hence, this research proposed a methodology that combines semantic web technologies with principles of semiotics and ergonomics, and empirical methods drawn from the emerging discipline of legal design, that was used to create and evaluate DaPIS, the Data Protection Icon Set meant to support the data subjects' navigation of privacy policies. The icon set is modelled on PrOnto, an ontological representation of the GDPR, and is organized around its core modules: personal data, roles and agents, processing operations, processing purposes, legal bases, and data subjects' rights. In combination with the description of a privacy policy in the legal standard XML Akoma Ntoso, such an approach makes the icons machine-readable and semi-automatically retrievable. Icons can thus serve as information markers in lengthy privacy statements and support the navigation of the text by the data subject.

Chapter 1

Introduction

April 10, 2018. *Facebook, Social Media Privacy, and the Use and Abuse of Data.* Hearing of the Senate Committee on the Judiciary and Senate Committee on Commerce, Science, and Transportation.

Senator Johnson: “Do you have any idea how many of your users actually read the terms of service, the privacy policy, the statement of rights and responsibilities? I mean, actually read it?”

Mark Zuckerberg: “Senator, I do not.”

Senator Johnson: “Would you imagine it’s a very small percentage?”

Mark Zuckerberg: “Senator, who reads the whole thing? I would imagine that probably *most people do not read the whole thing. But everyone has the opportunity to and consents to it*[1].”

[. . .]

Senator Kennedy [68]: “Here’s what everybody’s been trying to tell you today, and – and I say this gently. Your user agreement sucks. [. . .] The purpose of that user agreement is to cover Facebook’s rear end. *It’s not to inform your users about their rights*[2]. Now, you know that and I know that. I’m going to suggest to you that you go back home and rewrite it. And tell your $1,200 an hour lawyers, no disrespect. They’re good. But – but tell them you want it written in English and not in Swahili, *so the average*

[1] My emphasis

[2] My emphasis

American can understand it[3]. That would be a start".

This is an excerpt of the exchange between concerned US Senators and Mark Zuckerberg, the founder and CEO of Facebook, in the wake of the Cambridge Analytica scandal [127]. In the early months of 2018, the public debate around the processing of personal data reached an extraordinary level of media coverage, with members of the US and EU Parliaments, as well as common citizens, expressing concerns about how their personal data is gathered and used. In the words reported above, two striking elements emerge: firstly, Zuckerberg openly admits that he expects only a few people to read the Facebook's terms of service and privacy terms - despite the availability of such information and despite the lawful declaration of consent to the processing that Facebook's users are asked to make. In other words, Facebook's CEO places the complete responsibility of readership and comprehension if the terms on the users of his service: after all, Facebook has complied with the legal obligation of providing necessary information about how it will process its customers' personal data. Secondly, Senator Kennedy colorfully points out that the legal terms describing such processing are not written in a user-centered way, i.e. in a comprehensible manner that aims to effectively inform "the average American" about her rights, but they are rather aimed to discharge the company's liability.

This conversation revolves around some of the main debated points of the information paradigm, which is extremely topical in the modern age. Digital technology ushered in a new era for the disclosure and collection of enormous amounts of personal data, which nowadays can reveal intimate details and change lives in an unprecedented manner. Impressive technological advancements (e.g. Big Data, Internet of Things, Artificial Intelligence) are profoundly influencing human communication and societal values, causing major concerns. "In a flourishing online ecology, where individuals, communities, institutions, and corporations generate content, experiences, interactions, and services, the supreme currency is information, including in-

[3] My emphasis

formation about people" writes Nissembaum [211, p.33]; Acquisti adds: "[i]f this is the age of information, then privacy is the issue of our time" [16, p. 509].

The right to know which of our personal data is collected for which purposes is granted by the principle of transparency, one of the cornerstones of European Union's data protection law. Transparency, in turn, generates citizens' trust in digital services that ensures the prosperous growth of the digital market and the flourishing of their digital life. Normally, however, the general public is only aware of the top of the iceberg about the modalities and extent to which its personal information is gathered and processed [98].

Therefore, what are the reasons why data subjects disregard the legally-binding terms describing how their personal data will be used, but consent nevertheless to them? And which of these challenges can transparency solve? Chapter 2 provides a thorough literature review to answer those questions and, hence, sets the necessary conditions to begin the investigation carried out in the rest of the dissertation. The chapter starts from the origins of the information paradigm in EU data protection legislation: transparency has been historically deemed a necessary element to fight the asymmetry of power between the subjects providing personal data and the organizations collecting them. The chapter continues with an extensive analysis of the phenomenon of non-readership of privacy policies by presenting research from various disciplines that focuses on the discovery and examination of the elements that determine behaviors and decisions of data subjects relating to their privacy. Namely, studies of human-computer interaction have investigated how to make interfaces more usable to enhance people's privacy. Behavioral economics has provided evidence on individuals' actual decision-making processes, which is distant from that of rational decision-makers presumed by the law. Other research has examined the design of interfaces and services to respond to those cognitive biases that determine disadvantageous or risky privacy decisions. All of these studies agree on one point: the information paradigm as it is classically implemented is a failure, mainly because it is based on the idea of the data subject as a rational decision-maker (i.e. the *homo economicus*). Therefore, the hurdles that hinder data subjects from reading and understanding the privacy terms and from exercising free consent are consequently analyzed. Notwithstanding human bounded rationality, privacy-related communication can nevertheless benefit from many interventions to make it more human-centered and effective: a paradigm shift that revolutionizes traditional manners of conveying legal knowledge is needed.

Adopted by the European Parliament on April 14, 2016 and enforceable since May 25, 2018, the General Data Protection Regulation[4] defines the legal framework in which the present research is carried out. Taken as example all over the world, the Regulation introduces the obligation of transparency for the information addressed to data subjects with the goal of redressing information asymmetry. "The concept of transparency in the GDPR is user-centric rather than legalistic" specifies the Article 29 Working Party [30, p.5]: the communication shall be tailored to the characteristics of the intended audience and not merely be formally compliant, because "the quality, accessibility and comprehensibility of the information is as important as [its] actual content" [30, p.5]. Within this view, the innovation of traditional ways of presenting legal information assumes unprecedented relevance, while the potential of visual communication is explicitly acknowledged: information to data subjects can be provided in combination with machine-readable, standardised icons (Article 12.7) to give "in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing". The development and efficacy of such icons must be grounded in an "evidence-based approach" and motivated through "extensive research" [30, p.26].

This incentive to machine-readable, visual communication about data practices constitutes the conceptual core of the present dissertation and identifies two parallel and intertwined lines of research: the first direction is re-

[4] Hereafter: GDPR

lated to technologies for the management and (semi-)automated extraction of legal information, whilst the second is linked to the user-centered design of legal information. Chapter 2 investigates how the power of artificial intelligence and legal technologies can be harvested to make human legal content interpretable and processable by machines. It also identifies the features of machine-readable information that can be leveraged to semi-automatically display legal information with visual structure and icons in order to support reading and navigation. The information contained in legal documents must be conceived as the combination of actual textual content and additional machine-interpretable information that describes the structural and semantic meaning of the document.

The chapter outlines how standard mark-up languages (i.e. XMLs) for the legal sphere, such as Akoma Ntoso and LegalRuleML, can provide the syntax to assign machine-readable meaning to specific parts of the legal document: structure, semantics, and rules. But the original research mainly focuses on the design of an ontology that formally represents and organizes the data protection domain knolwedge. Such ontology is then necessary to provide meaning to the semantic tags and to allow automated reasoning on the text. The design of PrOnto, a GDPR-centered privacy and data protection ontology, is thus described at the end of Chapter 3. The conceptual modules in which the ontology is organized constitute the formalization for the development of the data protection icon set (DaPIS) that is described in Chapter 6: i) data (e.g. personal data); ii) agents and roles (e.g. data subject, controller); iii) data processing operations (e.g. anonymization, encryption); iv) processing purposes (e.g. marketing, profiling) and legal bases (e.g. contract, legal obligation); v) legal rules and deontic operators (e.g. data subjects' rights).

The use of technologies to make legal information more accessible to laypeople is one of the central assumptions of the emerging discipline of Legal Design, that is presented in Chapter 4 and can be defined as "the application of human-centered design to the world of law, to make legal systems and services more human-centered, usable, and satisfying" [141, Chap. 1]. Legal design prioritizes the point of view of all the 'users' of the law: not only that of lawyers and judges, but also citizens, businesses, etc. It is a lens through which to observe the *status quo* can be observed and redesigned in a 'human-centered' way: communication and interaction between individuals and the law are not designed for non-experts, even when they are explicitly addressed to them, as in the case of privacy policies and consent requests. Such recognized flaws in (privacy-related) communication prevent individuals to be informed about their rights and to exercise them.

The Chapter introduces a growing body of research that investigates the introduction of visual means in legal communication (i.e. legal visualizations), which is traditionally based on the written word. Among the many attested benefits, empirical research demonstrates that visual communication can unburden the cognitive load derived from reading and understanding complex legal information and make abstract concepts easier to grasp. Icons for data protection represent one but many of the examples concerning this vibrant research area that will be provided in the chapter. Another pillar of legal design is the reliance on empirical user research methods: users are involved in every phase of the design cycle from the discovery of their actual needs in the brainstorming phase, to the prototyping and evolution of solutions. A fundamental role assumes empirical research methods: iterative and measurable evaluation is active throughout the whole design cycle in order to evolve ideas meaningfully and measure the impact of the proposed solutions. The legal design research exits the exclusive realm of lawyers, actively seeks interdisciplinary collaborations, and even opens the doors to those individuals that will be actually impacted by the intervention: the data subjects. A central tool that will be described thoroughly is constituted by design patterns: replicable, systematized, and extensible solutions, as opposed to a jungle of bespoke different interventions. The most relevant and extensive research in this respect is about contract design patterns and privacy design patterns, that inspire the collection of legal design patterns for transparency and consent described in the following chapter.

Chapter 5 specifically addresses the prominent role of design of communication and interactions with technologies in the domain of data protection and privacy. Technology and evolution of mankind have proceeded hand in hand since the first man-made tools, while the rules governing societies, in terms of laws but also social norms, are profoundly impacted by technological advancements, with data protection representing a major example in this sense. The European stance on the complex interplay among technology, design, rules, and society is stated with crystalline evidence in Article 4 of the GDPR: "the processing of personal data should be *designed*[5] to serve mankind" and be based on the fundamental rights and values that shape our democratic societies. Chapter 5, thus, examines the role of design for the promotion of rules and values of European data protection law, with a focus on transparency?

Technology can be designed to achieve privacy-preserving outcomes by making it easy for data subjects to adopt privacy-protecting behaviors. With Article 25, data protection by design and by default are introduced as linchpins of data protection law: privacy requirements should be embedded into the design and architecture of any system, thus reflecting a pro-active attitude, while data subjects should not take any active action to protect their privacy, that should be guaranteed by default. There exists extensive research about Privacy Enhancing Technologies and privacy design patterns that implement the abstract principles of privacy by design - much less on technologies and patterns that aim to information transparency, as our analysis will show. User-centeredness, as one of the foundamental privacy by design principles, places the human being, rather than e.g. purely economic considerations, at the center of technology development, while providing her control over her data: privacy-preserving defaults, appropriate notice and empowering user-friendly options are examples of user-centric measures. Yet, design can also be used for the opposite aim: that of creating privacy-corrosive tech-

[5] My emphasis.

nologies and design patterns that purposely deceive users and lure them, for instance, into giving uninformed consent to processing. Uninformative and bad designed privacy policies and consent controls count as malicious (i.e. dark) design patterns.

Design choices can, thus, be used to promote privacy-conscious behaviors or, conversely, to facilitate privacy-eroding practices. Such considerations give rise to the next research question:

RQ1: Which legal design patterns can offer a solution to the problems of traditional disclosures and consent?

The last section of Chapter 5 analyses and systematizes possible solutions to those issues that hinder transparency of communication and informed consent emerged in Chapter 2, and proposes legal design patterns that can implement those solutions in practice. The patterns can act on three different levels (language, visualization, and interaction) and be combined to solve multiple problems. One of the visual patterns is represented by icon sets for privacy and data protection, a transparency solution that has seen some experimentation and that has been explicitly suggested by the European regulators. Our analysis better defines the role that such icons should assume in the specific context of privacy policies: they should act as information markers to ease content navigation of those lengthy texts and support strategic reading.

Chapter 6 is dedicated to the description of the development and evaluation of DaPIS, the Data Protection Icon Set at the center of this dissertation. The chapter focuses on the following research questions:

RQ2: What idiosyncratic features have icons with respect to other kinds of visualizations?

RQ3: Which challenges to creation and interpretation do icons present?

RQ4: What is the function and context of use of data protection icons?

Icons are different from other kinds of visual elements because, unlike diagrams or flowcharts, they are pictorial representations. Previous attempts to create icon sets for data protection are also critically examined in order to inform the design of DaPIS. A common misconception in the legal sphere is that icons are able to convey meanings universally. However, their ease of interpretation depends on several factors. At the individual level, dimensions such as familiarity, semantic distance, concreteness, and complexity of the icon must be explored, whereas also discriminability and coherence across the icon set are criteria that should be considered. Moreover, user's characteristics such as culture and level of experience with the represented concepts also influence the interpretation process. Icons for legal matters present an additional challenge compared to the majority of graphical symbols in use, that mainly depict concrete objects: legal icons mostly convey abstract meanings and represent unfamiliar notions. For this reason, not only their design but even their evaluation can be demanding, since arbitrariness or lack of familiarity cause low recognition rates at first exposures. From these considerations stem the last two research questions:

RQ5: How can icons for data protection be designed?

RQ6: According to which criteria can their effectiveness be evaluated?

Such research questions are more methodological than the previous ones, that have been answered mostly by a literature review and an analysis of the existing landscape. The computational approach based on the formalization of knowledge described in Chapter 3 is here coupled with the human-centered methods of legal design introduced in Chapter 4, both in the phase of creation and in the phase of evaluation of the icon set. Chapter 6 describes the methodology to generate DaPIS, that follows the circular design thinking methodology introduced in Chapter 4: ideas are gradually developed, in a constant generation of hypotheses that are iteratively tested and, thus, confirmed or rejected. The design was iterative: each development in the icon design was followed by an evaluation phase to determine legibility and comprehensibility of the symbols, whose results informed the subsequent (re)design of the icon set. Since some of the symbols are inherently arbitrary, thus their meaning can not be immediately evident, it was also researched whether the user could understand the reasons behind certain iconographical choices, i.e. if she could align her mental model with that of the icons' designers. DaPIS was created through a series of participatory design workshops, where interdisciplinary teams (mostly composed of lawyers and designers) confronted themselves with the tough challenge of creating small graphical symbols to convey complex and nuanced legal meanings.

The results of the evaluation will show that, unsurprisingly, the symbols that received higher scores represent concrete objects, familiar concepts or are based on familiar representations (e.g. the 'i' signifying information). Conversely, the concepts behind the icons that scored worst are vague, general, and abstract (e.g. the purpose of provision of the service). Chapter 6 ends with a thorough discussion about the results and suggests standardization of the icon set and education of EU data subjects to augment the ease of recognition of the DaPIS visual language, because some iconographical choices are inescapably arbitrary, while some others are uninformative, if it is the underlying referent to be unknown to the interpreters.

Chapter 7 provides directions for future research in order not to overlook any dimension of the evaluation of DaPIS: for example it delineates the necessity to determine a threshold of acceptability for the icons, by especially considering the intercultural nature of the EU residents and their varying levels of experience with data protection matters. This last chapter also describes thoroughly how to design and implement a comprehensive experiment of evaluation to test the effectiveness of the icons in contexts and the discriminability across the elements of the icon set, among other dimensions. In addition, some open questions that deserve further research are also introduced.

This chapter provides the research scenario that motivates the present dissertation: in indeed introduces the information paradigm in EU data protection law and the recognized hurdles it presents. Section 2.1 illustrates the rationale for the paradigm of transparency and control, that is rooted in the history of data protection law in the US and EU. Transparency is generally realized through the regulatory tool of mandated disclosures, while control over the flow of personal data is established through the instrument of informed consent. However, privacy disclosures usually fail to be informative and meaningful for data subjects, that do not rely on such information to make their privacy-related decisions, as much evidence shows.

Thus, in Section 2.2, the many limitations of the paradigm of transparency and control will be thoroughly analyzed, based on a non-exhaustive review of the growing body of literature that examines this phenomenon under multiple disciplinary perspectives. Essentially, the information paradigm is flawed because it assumes that individuals are rational decision-makers, while evidence shows that, in privacy as well as in other domains, human beings are rather subject to rules of thumbs and biases. This topic will be touched upon in Section 2.3. Any kind of architecture (e.g. information architecture, interface design) can leverage such human cognitive boundaries to design experiences that guide data subjects towards foreseeable choices: these can be beneficial or, conversely, detrimental to the welfare of the individual (or even of the society). Examples and implications of choice architecture employed in privacy and data protection will be provided in Section 2.4.

Section 2.5 discusses whether the information paradigm should be considered a failure and thus definitely abandoned, as some scholars suggest, considering the issues that it presents. However, we claim that interventions based on choice architecture can be put in place to design information and consent choices that are more easily understood by data subjects. The General Data Protection Regulation proposes, indeed, empirically informed solutions to address some of the concerns identified by research, that will be analyzed in Section 2.6. One of the main novelties is the introduction of the obligation of transparency, which focuses on the quality and accessibility of communication, rather than being confined to the content of communication. A major innovation is the acknowledgement of the support that visual elements, namely machine-readable icons, can provide in the navigation and comprehension of privacy policies. Such acknowledgement sets the foundations for the research described throughout this dissertation.

Chapter 2

The Information Paradigm

2.1 The Information Paradigm

2.1.1 Rationale

The information paradigm is a cornerstone of European data protection law. It is realized through the regulatory tool of mandated disclosures, which usually take the form of privacy policies[1] that provide details about how data controllers collect, use and protect data subjects' personal data, and about how data subjects can exercise their rights [205]. Essentially, such an approach informs users about the practices that will be carried out on their personal data before they access a certain service, therefore before the data processing starts. It is assumed that only through the analysis and understanding of complete and relevant information the data subject can exercise her right of control over the flux of her personal data: such information is considered the bedrock for the data subject to make the choice either to engage or disengage with the service (i.e. informed consent) [211].

The information paradigm postulates that the establishment of full trans-

[1] i.e. privacy notices or privacy statements or privacy terms. Throughout this dissertation, these terms will be used interchangeably to indicate those documents that describe the collection and processing operations carried out on an individual's personal data, although some other authors find relevant to differentiate among them, *see* e.g. [152].

parency facilitates individuals with their decisions about the permissible use of their data [215]. Many privacy regulations around the world move from these considerations: the notion of "privacy self-management" [266] implies that human beings are competent managers of their personal data and their privacy preferences [61]. This assumption is derived from neoclassical economic theories that consider the individual as *homo economicus*: a person that can access all relevant information at any time, that is able to understand that information, and that takes autonomous decisions in a fully rational manner, by carefully considering and comparing the trade-offs associated with the disclosure of her personal data [115].

The model of "notice and choice" is related to the notion of informational privacy [265] (*see* also [116]), which is, in turn, strictly entrenched in the notion of self-determination: individuals are free to choose optimally whether to disclose their personal data or to keep it for themselves. Such a view derives from the popular and traditional definition of privacy as control [211]: the right to privacy is the right to control information about oneself[2].

The notice and choice mechanism also derives from an economic perspective: if personal information is considered a digital currency in a commercial exchange (e.g. the provision of a certain service) in the competitive free market [123], then knowledge about the sellers' practices allows the customer to decide whether the price is appropriate and whether she wants to proceed with the transaction, as in any other purchasing decision[3]. In other words, individuals are free to determine what is an acceptable trade-off between the price for giving away their information and the services received in exchange [265].

However, such a transaction is fair only if the information in the hands

[2] Some classical definitions of privacy entrenched in the notion of control: Westin [294]: "privacy is the claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated to others"; Fried [113] defines privacy as "not simply an absence of information about us in the minds of others, rather it is the control we have over information about ourselves"; Elgesam [87]: "to have personal privacy is to have the ability to consent to the dissemination of personal information".

[3] Similar assumptions are intrinsic of EU consumer law, *see* e.g. [149], [123] and [96]

of the two parties is comparable, otherwise one party would be in an advantageous position with respect to the other [15]. Information asymmetries are standard situations in the privacy world, though, since the data controller (i.e. the organization collecting the data and deciding the purposes of processing) has more information concerning the data collection, its purposes and the practices of sharing and processing, compared to the data subject. Traditionally, the provision of information disclosures has been deemed an appropriate regulatory response to market failures that originate from asymmetric information [270]: mandated disclosures "protect the naive in dealing with the sophisticated" [40, p. 3]. Hence, privacy notices are expected to offer necessary and sufficient details to acquire an adequate understanding of what happens to one's data [265].

Finally, transparency about data collection and provision of meaningful choices have been recognized by several researches as linchpins to foster trust towards the service [197, 301, 207, 275] and thus encourage personal disclosure.

2.1.2 Origins

Gonzales [123] and Mantelero [196] have retraced the origins of the information paradigm in European data protection law. The idea dates to the 1960s, when the advent of the computer era caused the digitization and concentration of data in the hands of a few, mainly public, entities. The regulations of that period, thus, tackled the emerging power of computing that sparked the perception of loss of control over one's own personal data, combined with a general lack of awareness about this fact. As a response, transparency was increased and the individual's right to access the information held about her was introduced[4]. However, consent was still out of the picture since there was no space for self-determination and for economic exploitation of data. Along these lines were developed the principles sets forth

[4] *see* the Fair Information Practice Principles (FIPPS), later integrated in the 1980's OECD Privacy Guidelines

in 1981 by the Council of Europe's Convention 108 [78], which also added transparency requirements concerning the purposes and the place of data processing.

In the following decade, data revealed their economic value (e.g. marketing based on personal profiling) and data analysis arose. In this context, the Data Protection Directive [99] was written to pursue the economic interests related to the free flow of data and to protect the fundamental right to protection against unfair or unwanted exploitation of personal information. User's consent[5] became an instrument to exercise control over the personal data, but also to negotiate its economic value. The Data Protection Directive recognized the controller's obligation to inform the data subject as requirement for the fairness of processing: it set forth obligations about the information items to be provided to data subjects, regarded as a fundamental measure to promote transparency. This is when the notice and consent model was added to the existing paradigm of transparency and access. Technical complexity was still not deemed an obstacle to the understanding of the purposes of processing. Finally, Article 8 of the EU Charter of Fundamental Rights [94] provided a link between the right to information and the right to access and rectify one's own data.

The ePrivacy Directive [100] took into account these principles and required companies to obtain users' consent before placing and accessing data, like cookies, on digital devices. The same principles also form the cornerstones of the General Data Protection Regulation[6], the legislation introduced in 2016 and enforceable since May 2018 in all the European Member States. In Section 2.6, details about the GDPR's introduction of the transparency obligation will be further explored.

[5] For an analysis of the evolution of the meaning of consent in EU data protection legislation, *see* [61].

[6] Hereafter: GDPR

2.2 Issues of the Information Paradigm

Notwithstanding the historical, legal, and economic reasons behind the notice and choice framework as cornerstone of data protection legislation, research [202, 211, 190] and anecdotal evidence show that privacy policies are ineffective at informing individuals [260]: privacy policies are never or rather infrequently read, least of all by young people [193]. "[E]ven if they did [read the privacy terms], most individuals have difficulty fully comprehending what they actually agreed to and the risk they inherited by that consent" [147, pp. 1651-1652]. For example, a 2016 American study [214] revealed that 74% of the participants skipped the privacy information. Among those that did not, the average reading time was 73 seconds, indicating that the documents were not carefully examined. 98% of its participants missed a relevant clause in the terms stating that the user agreed to exchange his firstborn child with social networking access.

Self-reported data tell a similar story. A 2016 Eurobarometer [280] shows that only 20% of the respondents claimed to be always informed about the conditions for data collection and use, while around 40% declared to be sometimes informed. One third admitted to be rarely or never informed. Similarly, whereas under one fifth of Europeans maintained to read privacy statements fully, roughly half read them partially, while nearly a third admitted that it does not read them at all [279].

Often the readership or non-readership of privacy policies is presented as a free and rational choice of the data subject. The user is the only one to blame when she does not consider disclosures as tools to base her privacy decision-making: after all, the information is made available and she is explicitly asked to read and consent to the terms. "Consumers need to read the terms and conditions, and understand what it means when they tick a box on the screen" writes a member of the European Parliament following Mark Zuckerberg's appearance in front of European policymakers in the wake of the Cambridge Analytica scandal [269]. According to her, individuals need to take responsibility for the protection of their data and should therefore use the tools that data protection rules provides them "to empower themselves". After all, users have access to the privacy policies of website, apps and devices, so it is generally assumed that they consciously choose to disregard the terms at their own risk and to accept the consequences of such a choice [265]. Therefore, it is fundamental to investigate the reasons behind such a counterintuitive behavior.

A frequent answer tells that people do not care about their privacy [148], so they deserve the exclusive blame for such an attested behavior. However, recent Eurobarometer statistical data [280], gathered in view of the ePrivacy directive reform, tell a different story: confidentiality of communication and granting permission to access their data or monitor their activity is of great importance to vast majority of respondents. Similarly, [279] shows that most Europeans are concerned about the lack of control over their personal information and the monitoring of their activities.

This data shows that individuals do care about their privacy. But privacy notices are considered too long, unclear or too difficult to understand [279]. They have been described as "that dense, unreadable, boilerplate text[s] tucked away in some corner of practically every website and application on the Internet" [148, p. 64]. Given that "[p]rivacy policies are verbose, difficult to understand, take too long to read" [249], their effectiveness at informing about data practices, which should lead to mindful decision-making, is doubtful [172], while attributing the entire blame to the user is short-sighted.

2.2.1 Hurdles to Effective Privacy Communication

It is easier to blame the user (i.e. her incapacity or her ignorance) than to acknowledge that the design of privacy policies and consent interactions influence people's actions and privacy-related behaviors [148].

2.2.1.1 Language and Readability

A first hurdle to readership is represented by the readability of most privacy policies, which can be assessed through measures based on intrinsic characteristics of the texts (e.g. word length and sentence length), for instance the Flesch Reading Ease formula. Research [172] has demonstrated that many privacy notices go beyond the understanding of the average internet user[7], given that they require a high school or college education level to be understood [103], while in practice even college students have poor comprehension of the content of privacy policies [246]. Another study [214] found that, indeed, language difficulty was one of the main perceived obstacles to read privacy policies. To set this data in context, consider that the most prominent report on literacy levels in the EU [92] found out that one in five Europeans aged 16 to 65 have literacy difficulties, meaning that they can read at best simple texts, retrieve simple facts, or make simple inferences, while they are unable to understand longer or more complex texts, and interpret beyond what is explicitly stated in the text. This means that the literacy levels presumed by most privacy policies does not correspond to the actual literacy level of European population.

2.2.1.2 Document Length and Information Overload

It has been estimated that individuals with high school or college education would need between 29 e 32 minutes to read an average privacy policy of around 8000 words [214]. The same experiment shows that information overload depending from notices' length is the principal factor that disincentives users to read because it is perceived as a too much time-consuming activity, as also emerged earlier. The BBC's study mentioned above [60] highlighted that merely reading the privacy policies and terms and conditions of 15 popular websites would take around nine hours, with Spotify's legal terms summing up to 13000 words, almost as long as a Shakespeare's work. A 2017 study of 50000 English-speaking privacy policies [103] found that, on average, they are 1700 words (min. 30 - max. 70000) and 70 sentences (min. 1 - max. 4000) long.

[7] Such data confirm other research about terms and conditions [188], whose readability level is far beyond what a literate adult can possibly understand.

Moreover, too much information overwhelms users, which in turn do not react reasonably, but "skim, freeze or pick out information arbitrarily" [59]. The length of privacy policies can create information overload leading to increased stress, impaired judgment and a feeling of helplessness [153]. This evidence suggests that the information paradigm's central assumption, i.e. that providing more information gives more power to data subjects, is flawed.

The time needed to read and interpret the privacy policies is referred to as transaction cost: borrowed by economic theories, the term has come to identify any kind of expense required to perform a certain task [148]. Transaction cost plays a fundamental role in the decision of reading or, conversely, disregarding privacy policies [17, 305]: individuals carry out a cost-benefit analysis that considers time and effort required for reading and understanding and the expected advantages.

2.2.1.3 Vagueness of Terms

Vagueness and ambiguity are inescapable features of natural language and, therefore, of legal language [278, 88, 89], where it can be even fundamental to leave room for different interpretations [242]. By remaining vague, the privacy notice can be effortlessly adapted to the application of new regulations, while its flexibility can provide legal coverage for those actions (e.g. processing operations) that may occur in the future [248] .

However, in many cases, vague privacy terms are deliberately and intentionally used to actively deceive users and conceal privacy-invasive practices [44]. In some cases, the explicit goal of vague terms is to puzzle the reader about the intended meaning, thus ultimately hindering her right to be informed. A 2017 study on almost 500 apps and websites offering services or products in various domains [7] highlighted that usually only very high level descriptions are provided, instead of detailed and specific information about data practices. Organizations often fail to provide information about third party sharing, country of storage and form of access to one's own data. Vague terms commonly used in privacy policies have been extensively covered [242, 248], revealing recurring structures such as 'might' and 'certain', that leave readers puzzled about the actual occurrence of data processing: e.g. "We *may* collect information about you"; "we disclose *certain* personal data with third parties", etc.

2.2.1.4 Wrong audience and wrong goal

Similarly, an essential criticism that has been moved against traditional privacy communication [251], and that echoes a criticism addressed towards legal communication in general [43], concerns the envisaged audience of such communication: the information around the collection and processing of data is mainly drafted by lawyers for lawyers. Most of the time the provision of notices merely aims to fulfil the legal requirement of mandated disclosure, instead of effectively inform data subjects about the collection and processing of their personal data. In other words, this communication is not aimed to respond to the needs of the people that would most benefit from it: those that are impacted by the text, mostly non-lawyers [134]. For example, linguistic complexity cannot be grasped by individuals with low literacy levels. Within this view, privacy notices can be effective accountability mechanisms for companies and necessary auditing instruments for supervisory authorities or advocates, but "they are just not very good at notifying users" [148, p. 70].

2.2.1.5 An Impenetrable Wall of Text

Not only privacy policies do not consider the needs of their users, but they are also traditionally displayed as a "wall of text" [230] that is "impenetrable" to the human eye [97]. In late 2017, only a minority of the most trafficked sites on the web displayed some kind of visual mechanism that organizes the text in a more digestible manner, such as layered disclosures or navigation cues [134]. It is demonstrated that comprehension is hindered by visually undifferentiated text, while it is facilitated if the reader is guided towards those parts that are more relevant (i.e. attention hierarchy): information structure and information display play a key role in the support of comprehension and intellectual performance [230]. A straightforward manner to convey the document structure is to divide the text in digestible paragraphs, organize the content in the relevant sections, provide informative headings and differentiate font boldface, i.e. provide information architecture to undifferentiated text. Privacy policies should be considered as functional artefacts: within this view structural elements offer affordances, i.e. the functionality to easily find relevant information. However, traditionally legal communication "focuses only on the essence and precision of the rules, but not at all on the needs and abilities of the individuals tasked with understanding and acting upon such rules" [230, p. 342].

2.2.1.6 Lack of Comparability across Policies

As seen earlier, privacy policies are considered helpful tools to choose one service in a pool of similar services based on the particular data practices. Thus, it should be easy to consult and compare the different notices. Nevertheless, the lack of information architecture makes it hard to compare similar information about different services. It is unrealistic to expect that individuals read privacy policies word-by-word, from beginning to end. In fact, empirical research shows that people skim privacy policies to find answers to their questions or to compare two services' practices [198, 247]. This is why, structured formats (e.g. privacy nutrition labels) have been proposed [176, 177] to standardize and facilitate ease of comparison across privacy communication.

2.2.1.7 Complexity in the Big Data Era

An even more conspicuous problem about the informed consent framework derives from the nature of data processing operations on the contemporary digital environment. The actual possibility of transparency and self-determination in the present Big Data era is criticized [196], because information suffers from concentration in the hands of a few actors and is subject to highly complex processing: the knowledge asymmetry is aggravated [205]. The purposes of big data analytics and the possible outputs are difficult to describe, since hidden or unpredictable inferences and correlations are extracted from huge datasets - and not only on the individual, but increasingly on large communities. This complexity leads to vague descriptions in the notices about collection purposes, whilst it is questionable whether the exhaustive provision of details would impact the individual's possibility to full comprehension. This is, however, the same argument used against conciseness in privacy policies [152]: only well-written, long privacy statements can indeed explain those practices. Finally, data subjects are asked to derive a rational decision about single disclosures, whereas the magnitude of predictive analytics makes it objectively impossible to anticipate and evaluate the consequences of data disclosure [266].

2.2.1.8 Timing of Disclosure

Another critical dimension that does not allow the use of disclosures as decision-making tools is timing. Many legislations, as the GDPR, require data controllers to disclose information at the moment of collection of data, so that data subjects receive the necessary information for their decision on whether to engage with a certain service before they actually start to do so (for instance, before the subscription to an online platform or prior to the entering into a contract). Despite the necessity of information provision prior to consent, such an approach might be inefficient as for what concerns the exercise of control. Schaub et al. [260] discuss the importance of displaying a privacy notice at an appropriate time: excessive time distance between the moment of seeing a notice and the moment of making a privacy decision can modify user's perception of the notice and even neutralize the effects of the privacy protection. In other words, it is more efficient to provide punctual and limited necessary information to inform a privacy choice than to provide all the details prior to the use of the service. A good example is provided by the Facebook privacy check-up that alerts users if they are sharing a content publicly and gives them the possibility to change the privacy settings of the single post straightaway.

Furthermore, when the user is engaged in a certain activity, i.e. he has a certain primary task, the notice is experienced more as a nuisance [214] than as a useful privacy tool, causing individuals to deliberately ignore the notice. For instance, the cookie policy is provided at the moment of landing on a certain website to obtain consent prior to sending data to that website. However, in that moment, the primary task of a user is the navigation of the website, and not the acquisition of information about the website processing practices, so she immediately dismisses the cookie, by giving her consent without reading.

2.2.1.9 Lack of Familiarity and Expertise

The understanding of privacy disclosures also depends on how knowledgeable the data subject is about data processing and data protection. Most users, however, generally lack the necessary experience to understand and assess the consequences of their disclosure attitudes [266]. Individuals are not able to base their decisions on full information disclosures, and this effect is stronger the less the information is understandable and the less the individuals are experienced in the domain [215].

Indeed, experience and practice are key factors in decision-making, since they allow human beings to create patterns in their mind. Such patterns represent standard solutions to a problem and can be consulted immediately and intuitively (*see* also design patterns in Chapters 4 and 5). On the contrary, novices need to understand the facts disclosed in the notices with no or little experience, thus they must place the information in context, and consequently understand how to act upon it. In other words, the capacity of reaching good decisions is a skill which is developed over practice, rather than an exercise in analytical logic, and cannot be taught simply by providing (a lot of) information [40]. In this light, the time at which information is provided is also crucial because it situates such information in context.

One possible solution is the education of users, but it is estimated [215] that this would have only limited effects: indeed, it is unrealistic to expect people to become literate and knowledgeable in every area of life. Some critical authors even conclude that educational efforts usually end up in dismal results [40]. As will be discussed in the two final chapters, education to data protection is necessary in the modern digital world, but it does not constitute a panacea to everything.

2.2.1.10 Notice Fatigue

It is generally assumed that rational actors would analyze on a benefit-cost base whether to read or skip privacy policies (*see* Section 2.1). But, in fact, individuals receive an onslaught of notices and attention scarcity prevents them from carefully analyzing all this information [148]. Even if individuals actually engaged themselves in reading every privacy notice of the numerous services they use, this would result in an extremely time-consuming activity, estimated in an average of 244 hours per year, as a widely-cited study proved [198]. This assessment was made in 2008 and exclusively considered the websites' navigation of US citizens: considered the rapidly evolving pace of technologies and the rapidly increasing market of apps and devices, this estimation would arguably need to be revisited.

In other words, "[t]here is simply no way for users to weight all of the available pieces of information to get an accurate risk assessment for every personal disclosure they make" [148, p. 143] because there is just too many of them [71]. Considering the enormous number of entities with which the data subject interacts on a daily basis, both in the online and the offline world [266, 40, 148],"the burden of having to check all of this must be just too high" [71, p. 11].

Users choose not to inspect a service's privacy policy because they assume that the time cost (i.e. the transaction cost) would be not compensated by the benefits of reading [17]. In this case, choosing not to read is a deliberate choice. Directly linked to these issues is the user experience with privacy policies: if people have negative expectations because their past experience suggest them that privacy-related communication is boring, lengthy, overly complex, and unhelpful, then it will be hard to revert their assumptions and convince them to engage with it.

2.2.2 Hurdles to Effective Consent

2.2.2.1 Consent Fatigue

The consent mechanism shows limitations because of habituation effects, i.e. the desensitization of people to (too) many demands concerning their privacy [265]: "[t]he sheer number of choices that inundate users under a control regime is overwhelming to the point of futility" [148, p. 64]: under these conditions, choice is not an empowering mechanism, but it can easily become a burden that overwhelms and confuses the user. It is unavoidable to be susceptible to habituation effects: many activities of our digital life rely on automated, routinized gestures, such as accepting legal conditions without reading. Tellingly, consent fatigue (or "click fatigue" [26]) is correlated to notice fatigue: the number of requests in the digital context diminishes the supposedly warning effect of consent mechanisms. Thus, data subjects are often in the situation of waiving away their right through reflex clicks [205].

Therefore, on the one hand, the goal of asking consent is that of pausing the data subject in order to make her reflect about her privacy decisions. However, users live interruptions as 'nuisance factor' [115]: the process of obtaining their consent diverts them from their primary task. Given that too much information can be experienced as a burden instead of a valuable tool for decision-making, requiring consent for every transaction comes with a cost, that of quantity and complexity of choice: "control means constantly making choices, which is time-sapping and soul-sucking" [40, p.53].

2.2.2.2 Lack of choice

Other observations concern the meaningfulness of choice: in many cases, privacy notices offer information but lack any genuine choice about data practices. The tool of consent collides with the reality that most times data subjects are left with no choice but to agree [266]. In other words, in most situations they have no real bargaining power with the entity collecting their data and dictating the conditions. In the literature, this is called the "take-it-or-leave-it-approach": users need to accept privacy practices that they cannot negotiate in exchange of a certain service or are forced to go elsewhere [261]. As a consequence, data subjects might experience a feeling of helplessness. In such a case, the choice of not engaging with the reading activity is a rational choice.

Given the cost derived from not participating in the modern (online) social, commercial and financial life, it is also arguable to define this engagement as an individual's free choice to pay the informational price to access these services [211]. Where no comparable alternative is provided, then consent cannot be free [265].

2.3 *Homo sapiens* versus *Homo Economicus*

As the previous sections have illustrated, classical economic and legal theories consider data subjects as "the competent overseer of their privacy preferences" [61, p.465] or as precise calculators disposing of unlimited computational resources, applying this assumption derived by the economic theories to privacy decision-making [15]. However, many legal scholars have accepted the reality that, even if individuals claim that they worry about their privacy and that they are conscious about their rights, they do not act accordingly: an inconsistent attitude called privacy paradox [205]. For instance, even if they could read privacy notices to acquire that information, they do not. Furthermore, provision of accurate information about risks inherited by privacy disclosures does not change users' behaviours [205].

The reasons for such a discrepancy lie in the evidence that the information paradigm is an abstract principle that does not correspond to actual human behavior. Many assumptions about how people make decisions are simply false [266] and disregard "the real-world experience of users as biddable and bewildered" [61, p. 465].

Behavioral economics research, which studies "how individual, social, cognitive, and emotional biases influence economic decisions" [17, p. 368], has exposed many hurdles in decision-making that are not only related to the analysis of privacy disclosures and consequent consent, but also related to privacy and security in general [15] and in many other domains of life (such as healthcare, nutrition, finance, or environment). Such studies have made evident how biases (*see* below Section 2.3.2) and heuristics (*see* below Section 2.3.1) influence human behaviors and diverge from classical assumptions of economic theory: namely, the abstract model of *homo economicus* does not correspond to the actual behavior of *homo sapiens*.

Even if users did access and understand exhaustively all the information that could inform their privacy-related behaviors, they would be unable to consider all the consequences of data disclosure, due to human innate bounded rationality [17], namely their limited cognitive resources and behavioural biases. For example humans tend to base their decision-making on heuristics rather than on rational deliberation.

Empirical evidence shows that increasing the quantity of information presented to users does not help them with their decisions [215]. The information paradigm should also consider the quality of information under many perspectives (e.g. in terms of readability or presentation formats), as it will be several times asserted during this dissertation. This is why some scholars call for "empirically informed approaches" [270]: setting rules that take into account human bounded rationality and actual human behaviors. Disclosure policies should rely on an empirical understanding of how individuals actually make sense of information: “disclosures requirements should be designed for *homo sapiens*, not *homo economicus* ” [270, 1369]. Clarity and simplicity are key notions in this sense: disclosures should be “concrete, straightforward, simple, meaningful, timely, and salient” to be useful and helpful [270], and not merely legally complaint. Moreover, disclosures should be also contextualized in time and space [260], i.e. that they must be presented at the time of decisions, as it has been argued earlier.

Ameliorating the usability and comprehensibility of privacy policies can enhance individual’s understanding of how her data is used and thus reduce information asymmetries between data subjects and controllers. However, increasing transparency does not have effects on two further levels of problems. The first one is the human innate bounded rationality: to face complexity of existence, human beings rely on cognitive shortcuts, or heuristics, that differ from rational decision-making methods. Especially in the complex data ecosystem, it becomes nearly impossible to evaluate attentively the likelihood of privacy risks and the costs of privacy disclosures. Secondly, even if human beings had sufficient cognitive capability to process that information, they suffer from cognitive biases, i.e. systematic deviations from behaviors postulated by rational economic theory [17]. “Given that privacy’s tangible and intangible consequences are often difficult to estimate, numerous heuristics and biases can influence and distort the way individuals value data protection and act on privacy concerns. A growing body of empirical research has started highlighting the role of such systematic inconsistencies in privacy decision making” [18], which will be briefly analyzed in the next sections.

2.3.1 Heuristics

Also known as rules of thumb, heuristics refer to shortcuts in decision-making that are regularly employed when the innate bounded rationality of human beings impair their ability to analyze all the possible options or outcomes of a certain action. In privacy decision-making, this means that individuals do not (or cannot) rationally consider the trade-offs between risks and rewards relating to the disclosure and protection of personal information. They rather base their decisions on perfunctory judgment, e.g. on the price of a certain service rather than on its privacy protection. For instance, research shows that individuals simplify their reasoning about the likelihood of an event (such as a privacy harm) leaning on events that are more or less readily accessible in their memory: the less available the more the risk will be underestimated [270]. This fallacy in rational thinking is known as the availability heuristics. Another rule of thumb that drive users to sub-optimal decisions is hyperbolic discounting, which can impact considerations about advantages and disadvantages of a certain choice: long-term events (e.g. privacy harms) are perceived as more distant and thus less important than short-term benefits (e.g. the free access to a service).

2.3.2 Biases

Biases are “systematic, therefore predictable, deviations from rational choice theory” [15, p. 44:4] and are independent from the complexity of a certain choice. Among the many biases to which privacy disclosure behaviors are susceptible (for a thorough analysis, *see* [15]), for the current investigation will exclusively refer to those relevant to ameliorate the efficacy of privacy disclosures.

2.3.2.1 Framing

Increasing or decreasing the salience of a certain piece of information over others can trigger different behaviours: for example, people seem more aware of privacy risks if these are presented through an alert or an image instead of a text [204]. The so-called “framing effect”, i.e. the presentation modality of a piece of information, plays a role also in the choice between concrete benefits and abstract, future risks: this is why, the way privacy policies present advantages and disadvantages of data disclosure can have a profound impact on the individual’s choices [61, 64]. Framing the same information as more or less protective, for example, can influence the disclosure behaviors of users even when the objective risk related to the disclosure is not altered [18, 15]. Namely, emphasizing privacy protection in the notice lured the users into more disclosure.

The way in which information is communicated *per se* is never neutral and its comprehension alone does not counterbalance the effects of deceptive framing (*see* also Section 5.3). Organizations that base their business model on personal data collection can arguably frame the information they present to users in such a way that consent is easily given.

2.3.2.2 Inertia and Status Quo

A further relevant bias in this context is the inertia bias. Only rarely individuals challenge the status quo, because people tend to stick with the default option that is provided to them. For instance, the majority of people only rarely change their default privacy settings on a device or a platform and they tend not to uncheck pre-ticked boxes, hence they inadvertently give consent to data practices. As it was recalled earlier, it is not easy to form an opinion on rather complex or unfamiliar topics, thus many people tend to believe that there was a well-grounded reason for a specific default privacy setting.

However, default rules can also have beneficial effects: in principle they spare individuals the burden of time-consuming choices, but preserve their freedom of choice [270]. However, they raise the question about which default rule should be the preferred one. One approach favors the default rule that the majority would select if adequately informed. Another reasonable criterion considers the default rule that ensures automatic compliance with the law. However, default rules can be badly chosen or misused, and even considered a sort of manipulation.

Two alternatives to default opt-ins can be envisioned. The first is represented by default opt-outs, which presumes user's activity to signify consent. The second solution is represented by active choice, especially when the target group is so diverse that it can be hard to determine which would be the preferred solution for the majority. Nevertheless, active choice also suffer from disadvantages: when human beings lack knowledge or experience in unfamiliar or complex situations, active choosing might increase the costs of decisions and place an heavy burden on the subject (*see* e.g. consent fatigue). When choices multiply, then this strategy might outweigh any advantage, because the number of choices might simply become too high to be easily managed.

2.4 Nudges

The debate introduced in the last pages must be understood and considered in the wider frame of the discussion about paternalistic approaches versus libertarian approaches. On one side, coercive regulations, in the form of mandates of bans, force individuals or organizations to act in a determined way. On the opposite side, self-regulatory solutions proclaim the absolute self-determination of individuals and assume their ability to make choices in their best interest. The examples reported above show how both approaches might fail, whereas a growing number of scholars argue that soft [15] or libertarian paternalistic [272] interventions might guide users towards behaviors that promote personal welfare. Sunstein speaks of nudges, i.e. "liberty-preserving approaches that steer people in particular directions, but that also allow them to go their own way" [271, p.583]. Nudges are changes in choice architecture that are intended to encourage certain behaviours [205]. The goal of a nudge is that of reducing individuals' burdens when they attempt to achieve their goals and have been used more and more in the private and public sector because of their effectiveness.

"It is pointless to object to choice architecture or nudging as such" affirms Sunstein because "choice architecture is inevitable" [272, p. 6 and 44]: anything, from department stores to websites, has an architecture that is expressly designed to encourage or discourage certain choices. For example, the position of fresh vegetables or sweets in a supermarket exerts predictable and observable effects on customers, i.e. it steers them towards more or less healthy purchases. Similarly, the way apps, websites and social platforms are designed is made to direct their users towards certain behaviors and choices. If these are beneficial or not for the individual depends from the intentions behind the design architecture. On a positive side, nudges for privacy and data protection can complement the traditional paradigm of notice and choice in order to cover for its inefficacy (e.g. *see* the results of [205] and Section 5.2). On the negative side, entities with unethical goals can use them to lure users away from privacy-conscious behaviours [205, 9] (*see* Section 5.3). In the following some examples of privacy nudges that are relevant for the discussion presented in this dissertation will be explored.

2.4.1 Examples of Nudges

2.4.1.1 Relevant Information and Choice

Default rules (*see* Section 2.3.2.2), like mandated opt-ins, and the salience of relevant features are nudges: they are meant to provide relevant information to people so that they can make better decisions for themselves. They try to compensate for the fact that "[a]ttention is a scarce resource" [272, p. 8]. To overcome information and power asymmetries, a typical nudge is the increase of transparency and choices. However, research shows that these interventions do not always reach the desired outcome [18] and might even be detrimental to the individual [41]. Indeed, an excessive amount of information and choice might make the individual feel overwhelmed (*see* Section 2.2.1.2), whereas well-thought design that present relevant and concise information at the right time might achieve the desired goals [15]: for instance, the specificity about processing purposes might increase disclosure.

2.4.1.2 Framing

Framing can also act as nudge, for instance when users are mandated to decide upon benefits and risks associated with personal data disclosures.

Since the same content can be presented in different manners, the selection and salience of specific aspects significantly affects how people perceive it: "frames enhance the probability that receivers will interpret information in a certain way, discern a particular meaning, and process it accordingly" [148, p. 38-39]. Saliency makes people focus about certain features and disregard others: the power of this technique also lies in the omission. Framing and saliency can be used to present the benefits of sharing personal data (*see* also [30]) but hide potential risks. For example, Section 5.3 will provide an example about how Facebook has maneuvered its users to consent to face recognition, by framing this choice as a security measure. The effects of salience, therefore, depends on the intentions of the entities that provide information and seek for consent: if their interests are malevolent, they can easily shape and frame the interaction with data subjects in such a way that their consent is easy to extort.

2.4.1.3 Availability

The ease of availability of one option over another, even if potentially equal, can make it harder for users to safeguard their privacy. An example is represented by promotional emails: whereas it is easy to subscribe, the unsubscribe option is typically placed at the bottom of the message, in small fonts and bland colors. Buttons can also be leveraged to nudge users towards data disclosure. For instance, active (i.e. colored) buttons are more attractive than non-active (i.e. greyed-out) buttons: a fact that invites users to click them. Even the position of buttons on the right of a desktop's window or closer to the thumb can imply forward movements and can make a certain movement more natural than another. When a button corresponds to consent, a prototypical situation in the digital world, it is not hard to understand the implications of such design.

2.4.1.4 Structure

When a user is confronted with complex decisions where she must evaluate and decide among many alternatives, then structure, like structural layout, is a mechanism that invites and eases the comparison, whereas the lack of structure triggers the opposite effect.

2.4.2 Every Choice is a Nudge

One argument against nudges concerns their supposed intrusion on autonomy: for example, the provision of privacy-friendly default options, as opposed to a more, seemingly autonomous, active choice. But there is no autonomy without informed choices, and many nudges attempt to make choices more informed. Not only: when they help correct inherent biases, nudges might actually promote individuals' autonomy. However, autonomy cannot be identified *tout court* with active choices in every context. On the contrary: increasing the number of choices (e.g. the number of consents) actually reduces autonomy because individuals cannot focus on those issues that in their opinion deserve that attention.

Nevertheless, default rules might intrude on autonomy if they do not adhere to people's preferences. Acquisti et al [15] share this view: fears about nudges are overstated, although nudging does come with ethical considerations and implications. Nevertheless, "every design choice inescapably influences the user in some way" [15, p. 44:27], whether it was intentionally designed to affect individuals' behaviours or not. In the design of a user interface, every choice architecture is explicitly made to guide people's behavior. Even minor changes to the decision environment influence choices, which typically goes unnoticed by the decision maker.

"Many nudges, and many changes in choice architecture, are not merely permissible on ethical grounds; they are actually required." ends Sunstein [272, p. 50]. The question is therefore not whether nudging is ethical *per se*, but rather if a specific nudge is ethical in specific contexts [15], i.e. if it has licit or illicit goals. For example, the same nudge can be ethical for some user, but not for others. Nudges can be used to align behaviors with stated preferences, while users' choices must be respected - for example, they must be allowed to adjust the nudge's settings. “As individuals within a population differ in terms of preferences, awareness, knowledge, but also personality traits or susceptibility to biases and heuristics, individually tailored nudges may result in more effective interventions” [15, p. 44:31]. Customization to individual preferences will also be discussed at the end of this dissertation.

2.5 The End of the Information Paradigm?

As the previous examples have shown, every choice architecture can influence users and steer them towards desired behaviours, which can be licit or illicit. Earlier, it was illustrated how the lack of information structure dissuades data subjects from reading privacy policies, while the framing of choices can encourage personal data sharing. Solutions that enhance transparency are criticized because the simplification of language or similar interventions alone will not solve the non-reading problem [151], as much evidence suggests (*see* e.g. [40]).

For all these reasons, the tool of mandated disclosure, on which informed consent is based, has attracted fierce criticism. Some scholars have come to the conclusion that the notice and choice mechanism has failed [211], not only in the privacy domain, but in any context where it is required [40], and have therefore called for a complete abandon of mandated disclosure as a regulatory tool [41].

The analysis of human innate bounds has brought some authors to conclude that consent by an individual cannot be regarded as a rational articulation of the individual's view: “any legal rule that treats consent as the product of a rational thought process is potentially open to question” [61, p. 471-472]. Given human irrationality, some scholars [61, 204, 266] even suggest that data protection law cannot continue to be based on the notion of informed consent.

It is well outside the scope of this dissertation to discuss the role of a more or less paternalistic regulation regarding data protection[8] or to propose alternative models to mandated disclosures.[9] However, it was deemed necessary to provide a broad picture of the complexity of the phenomenon of non-readership of privacy disclosures and to provide evidence that it does not exclusively depend on the individual's good will.

Although some authors call for the end of the paradigm of notice and choice, some other insist on the safeguard of the role of information provision and on the possibility of exercising informed consent in the modern data protection framework. Yet, the analyses illustrated above aimed to clarify that there are many problems that call for action. All in all, initiatives towards more education, more meaningful notices and more freedom of choice are laudable and important [266].

There are two main arguments that prevent the abandon of privacy disclosures. Firstly, although much evidence shows that the assumption of a rational decision-maker is far from realistic, still two thirds of all legislation is based on the information paradigm [215]. As it will be illustrated in the next section, also the General Data Protection Regulation sets as linchpin of its principles transparency and consent. Nevertheless, the Regulation also specifically describes how this mechanism must (and must not) put in place.

This observation introduces the second argument in defence of mandated disclosures. Some of the known hurdles to effective privacy communication can be partially solved by providing visual, simple, and user-friendly instruments [205] that challenge the status quo of traditional notices. For example, structure can counterweight the non-comparability of privacy policies and can offer affordances against walls of text, while framing can be used to give relevance to certain information over other. Opt-ins and active choices can counteract the inertia bias. Although this will not constitute the ultimate solution (for a thorough discussion, refer to [203]), privacy policies and con-

[8] E.g. [266] and [148] discuss the topic at length

[9] E.g. [40] propose peers or expert advice and rating systems.

sent mechanisms can be better designed, as will be exhaustively illustrated in Chapter 4 and 5.

It should also be considered that well designed transparency mechanisms will not have the same effect on any data subject. Well-designed disclosures are critical for consumers to make good choices, but only if they take into account the diversity between experienced users that might benefit more detailed information, and novices, that might need concise disclosures. Given that also context is key [148], an effective solution is represented by a careful selection of the timing of privacy notice display [260]. Such attention to timing, in addition, offers the considerable advantage of splitting lengthy notices in limited spans of information that are relevant to a specific task. Encouraging results from a growing body of literature on good legal information design (*see* Section 4.3), supported by the regulatory provision that will be described in the next paragraphs, suggest that it is worth attempting to change the *status quo* for a better implementation of the principle of transparency.

In conclusion, "transparency and choice may not be sufficient conditions for privacy protection." [18, p. 2], but only one of the possible tools to address individuals' concerns about privacy. In the following, it will be explored to what extent the GDPR considers the criticism moved against the traditional notice and choice framework and the solutions it offers.

2.6 Behavioral Insights in the GDPR

The GDPR, applicable in every EU member state since May 25, 2018, introduces many innovations compared to the previous legal framework: "[t]he broad aim of the GDPR is to encourage the emergence in real-world settings the active and empowered users that previous approaches have only presumed" [61, p.467]. Indeed, lessons learned about the psychological processes underlying behavior can be used to actually aid that behavior, by designing tools and policies that enhance choice, without restricting it [15,

p. 44:32].

The GDPR can be considered, at least in some of its provisions, an example of design-based regulation that embeds regulatory standards into the design of the system being regulated [302] (*see* Chapter 5), "i.e. to create an architecture for human behaviour that 'hardwires' in the preferred behavioural patterns" [189, p.58]. The provisions introduced in the Regulation that consider behavioral insights are outlined and analyzed in the following.

2.6.1 Empirically Informed Provisions in the GDPR

2.6.1.1 Framing

According to the GDPR Recital 39, data subjects should generally be "made aware of the risks, rules, safeguards and rights in relation to the processing of personal data". This means, for instance, that they should be adequately warned of any processing where collection of data is not obvious. Articles 13 also states that the consequences on the data subject derived from the failure of provision of data and of automated decision-making must be clearly spelled out, probably in an attempt to transform privacy policies in decision-making tools. In this way, controllers are obliged to clarify negative or possibly risky consequences of data sharing, instead of exclusively focusing on the positive outcomes of such practices. Concretely, this provision suggests that the salience of potentially harmful privacy terms and deviations from reasonable data subject's expectations can be enhanced [149], for example through the prominent display of graphical symbols, similarly to warning symbols. However, it is hard to determine what those reasonable expectations are. Moreover, the highlighting of risky data practices would go against data controllers' interests. As it will be thoroughly explained in Section 6.1.4.2, icons that upfront and clearly show unfair practices have been proposed (*see* [281, 102]), but it is hard to determine the incentives that would prod data controllers to display them.

2.6.1.2 Data Protection by Default

Perhaps the most striking example of behaviorally informed provision in the GDPR consists in the principles of Data Protection by Design and by Default, set forth in Article 25. As explained earlier, privacy-friendly defaults leverage, while at the same time counteract, the inertia bias. This topic will be explored thoroughly in Section 5.1.

2.6.1.3 Default Opt-Ins or Active Choices

Under the GDPR, consent must be free, specific and, especially, informed. Informed consent means that, before users agree on certain data practices, they shall be given notice in an available, visible and easily understandable way [31]. Quality of the information is deemed necessary to provide informed consent, that must be presented in an "intelligible and easily accessible form, using clear and plain language" (art. 7).

The GDPR strengthens these requirements by introducing explicit consent for any kind of data processing: consent cannot be inferred by the inactivity of the user, on the contrary it must be expressed "either by a statement or a clear affirmative action" (art. 42). This model presumes activity of users: they must opt-in to consent to the processing instead of opting-out from the default presumption of consent [61]. The data subject is thus lead to a dynamical engagement, whereas any ambiguity that might be signified by her passivity is avoided. Active choices, as shown earlier, counterbalance and, at the same time, leverage the inertia bias. They take into consideration the criticism moved along the years to the empirical implementation of the consent mechanism, traditionally based on default opt-ins. The consent tool as legitimate ground for processing has been abused, especially in the digital context, and has been considered ineffective at the very least. Hitherto, pre-ticked box or switched on toggle bars have been the norm and have tricked users into inadvertently giving consent by exploiting the status quo bias.

Given the relevance of the digital transactions, the GDPR provides quite precise indications about the choices related to interface design: “silence, pre-ticked boxes or inactivity” (Recital 32) are not considered lawful, since in these cases consent is presumed by default. Conversely, the thick of a box or the click of an icon are deemed acceptable because they signify the data subject’s unambiguous expression of her will through an active choice. Furthermore, innovative ways to signify consent are listed by the Article 29 Working Party[10] “Swiping on a screen, waiving in front of a smart camera, turning a smartphone around clockwise, or in a figure eight motion may be options to indicate agreement”, whereas “[s]crolling down or swiping through terms and conditions which include declarations of consent [...] will not satisfy the requirements of a clear affirmative action” [26, p.17]. The exact instructions provided to manufacturers, whose creativity is also called into play to solve the thorny issue of unambiguous consent (that can be easily become a nuisance if considered at scale, as discussed earlier), reveal regulators’ considerations about how interface and interaction design is crucial to signal to users that they are taking a legally-binding action.

2.6.1.4 Ease of Consent Withdrawal

Another example of the lessons taken from empirical research in the GDPR is represented by the specific provision about consent withdrawal: “[i]t shall be as easy to withdraw as to give consent” (Article 7.3). For instance swiping a bar in one sense or the other is considered lawful, whereas requiring undue effort would rely on user’s inertia bias, thus it shall be unlawful. Such is the case of giving consent to marketing communication through a mouse-click, but being forced to call a call-center to opt-out. It is questionable if also the example reported earlier about the subscribe and unsubscribe option would be lawful under the GDPR.

[10] Hereafter: WP 29

2.6.1.5 Against Notice and Consent Fatigue

Habituation effects can be fought, primarily, by determining whether the display of a notice at a certain time or in a certain context is necessary: after a few repetitions, it is well known that the notice goes completely unnoticed [260]. Although forcing interaction with the notice can reduce habituation effects, it can also result in excessive nuisance, causing the user to stop using the service. This is why, for example, the indication of consent can be expressed through cookies registering users' preferences. After the GDPR came into effect, many websites updated their cookie permissions, some of which offering detailed, purpose-dependent consent options to comply with the "specific consent" obligations. However, in this way the user has to restate her choices for every website or family of websites she visits for the first time. It is however foreseeable that such preferences will be managed via browser settings, as the proposal of the ePrivacy regulation suggests, or by automated privacy assistants [187].

2.7 Transparency

Earlier in this chapter (*see* Section 2.1), the origins of the information paradigm have been traced. In the following, the most recent evolution of transparency and the reasons why in the GDPR it becomes a funding principle and an obligation will be explored. This analysis will pave the way to an investigation into the empirically-informed provisions on transparency, that attempt to solve the issues outlined at the beginning of this chapter.

2.7.1 Origins of the Transparency Obligation

Although the introduction of the duty to inform in the Directive 95/46/EC marks a milestone, data protection authorities realized early that such a duty was more than often put into practice in an incorrect manner, resulting in very long disclosures that contained technical jargon and legalese. A 2009 study sponsored by the UK's ICO [251] highlighted that the obligation to inform was implemented poorly: "[p]rivacy policies are written by lawyers, for lawyers, and appear to serve little useful purpose for the data subject due to their length, complexity and extensive use of legal terminology." [p. 29]. The study also instils doubts about the extent to which privacy policies can be considered helpful tools to enforce data subjects' rights (doubts echoed in [93]). Therefore, it suggests that these documents might be more useful as data protection authorities' enforcement tools to check a company's self-reported privacy commitments by or in case of law infringement. This is why supervisory authorities called for more readable formats, such as multi-layered notices [22].

In 2010, the European Commission presented its approach to the future of data protection regime [93], introducing specifications on easy-to-understand, plain and clear language and suggested the introduction of a general principle of transparency, backed then by the EDPS who asked for an explicit principle of transparency [95]. With this document, transparency is transformed from a formal requirement into a more concrete indication of the manner of information provision.

2.7.2 Transparency under the GDPR

The document paved the way to the introduction of the transparency obligation in the GDPR, which is ascribed among the principles of lawfulness and fairness of processing (Article 5), thus affirming transparency as one of the cornerstones of EU data protection law. Although it was already alluded in Recital 38 of the Directive 95/46/EC, it is only with the General Data Protection Regulation that transparency becomes an overarching obligation [30] that applies to any communication addressed to data subjects, such as privacy notices, consent agreements, and data breach notifications.

Similarly to consumer law's transparency provisions[11] [149], data protec-

[11] *see* e.g. Article 5(1), 6(1) and 8 of the Consumer Rights Directive and Article 5 of the Unfair Terms Directive. On the wake of Facebook's CEO hearing at the US Senate, a

tion provisions focus on content and language. This information must be disclosed in a “concise, transparent, intelligible and easily accessible form, using clear and plain language, in particular for any information addressed specifically to a child” (art.12 [101]).

Transparency is not defined in the GDPR, but recital 39 provides indication about the nature of this obligation and about its effects: “[i]t should be transparent to natural persons that personal data concerning them are collected, used, consulted or otherwise processed and to what extent the personal data are or will be processed. The principle of transparency requires that any information and communication relating to the processing of those personal data be easily accessible and easy to understand, and that clear and plain language be used. That principle concerns, in particular, information to the data subjects on the identity of the controller and the purposes of the processing and further information to ensure fair and transparent processing in respect of the natural persons concerned and their right to obtain confirmation and communication of personal data concerning them which are being processed”.

The key provisions that concern transparency are to be found in Chapter III, in the articles that apply to data subjects. Article 12 defines general rules that apply to any information that must be provided to data subjects: not only the information about data collection and processing, but also about their rights and in case of data breach notifications. Much emphasis is placed on the quality of communication, that must be “concise, transparent, intelligible and easily accessible” and expressed in “clear and plain language”, while particular attention must be devoted to children.

bipartisan initiative proposed the “Social Media Privacy Protection and Consumer Rights Act’ that includes specific provisions about the terms of service which should be, among the others, “*of reasonable length*” and with language that is “clear, concise, and *well-organized*, and follows other best practices appropriate to the subject and *intended audience*” (my emphasis).

2.7.3 Quality of the information

The information paradigm traditionally revolves around the quantity of information, leaving out any indication about its quality [215]. However, under the GDPR, a shift seems to be occurring: even the quality, accessibility, and comprehensibility of privacy communication assume an unprecedented importance to demonstrate compliance with the principle of transparency [30]. "Compared to the DPD, the GDPR now includes rules on how the information must be presented to data subjects and not only which information should be presented" [203, p. 509].

A substantial innovation is constituted by the fact that the principle of transparency can be effected not only through verbal means but also through visualisation tools [30]. Under this light, the GDPR suggests to provide information in combination with machine-readable, standardised icons (Article 12.7) to give "in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing"[12]. This approach is aimed at reducing excessive amounts of written information [30]. Although eventually it will be the role of the European Commission to adopt delegated acts to give directions on the creation of these icons, the need of expert advice is emphasized in Recital 166 GDPR and in the dedicated Guidelines on Transparency by the WP29, which encourages an "evidence-based approach" and "extensive research" [30, p.26] to inform the development and application and determine the efficacy of icons in this context. The research described in Chapter 6 intends to contribute to the scientific debate around the icons.

2.7.4 User-centeredness

The expected audience of communication assumes fundamental importance: disclosures cannot be addressed to an abstract, average data subject, but should be rather tailored to specific users in a specific context. For instance, a teenager uploading pictures on Instagram has probably different

[12] Similarly, Article 8 of the proposed ePrivacy Regulation

cognitive needs and past experiences than an adult opening a bank account. This is why particular attention is also devoted to the language that must be used with children, as recalled earlier.

In fact, "[t]he concept of transparency in the GDPR is user-centric rather than legalistic" specifies the WP29 [30, p.5]. In other words, the intended audience of privacy communication and the characteristics of human cognition must be taken into account to provide understandable information. Instead of mandating what data subjects should do, user-centric transparency addresses actual users' needs and actual users' behaviors. Empirical research, for instance user research, is necessary to demonstrate the intelligibility and effectiveness of information notice. Documentation about how these studies are conducted and about the results can support the controller in complying with its accountability obligations. It is noteworthy to mention the specificity of vocabulary used in the Guidelines, which resonate usability research and design: words like "user-centric", "user testing", "user experience" are mentioned throughout the whole document. Data subjects are in the first place seen as users of a website, app, service or product that collects and processes their data. Even the type of device, user interface and circumstances of provision, such as timing, are considered crucial elements for the appropriateness of information provided.

Multi-layered notices represent an operative way to address different audiences: a layout of three combined different layers (short notice, condensed notice, and full notice) can improve readability and comprehensibility with respect to the device and time limitations, while offering compliance in its totality. Other attempts offer plain language next to jargon-filled versions of the same policy. Whereas only the latter has legally-binding value, the first is a translation into plain language terms that, in principle, satisfies most informative needs of the prototypical reader: not a lawyer, but a data subject that has only as few notions on data protection.

In order to fight information fatigue, the use of layered notices in an online context is now recommended to provide concise information [30], as required in Article 12.1. This means using a structured, navigable layout, where the data subject can easily find the information she is looking for and jump to the relevant section through direct links. A standardized layout can also help data subjects to compare different privacy policies (*see* e.g. [4]). Such structured labelling approach [40] can be integrated with icons that signpost the various items of information, similarly to the standard format provided by the Directive on Consumer Rights [227] [149].

Busch [55] recognizes in the layered notices and the "push/pull notices" described in the Guidelines a potential opening to customization of privacy disclosures. Indeed, the WP29 explicitly suggests the adoption of transparency tools that are able to display "tailored information to the individual data subject" [30, p.17]. In Chapter 7, customization of content display will be discussed at length.

2.7.5 Relevant Information

To be meaningful and draw users' attention, privacy policies should contain relevant information for them. It would be therefore important to identify the possible audiences of a privacy notice, hypothesize the unexpected practices for each audience and give more salience to them [260]. For instance, the user of a torch app does not expect it to share its location data with third parties. However, it is also important to explain the reasons why a certain unexpected practice is active and what are the benefits of it, while not hiding privacy risks. If possible, there should be actual empirical evidence, e.g. surveys or experiments, supporting what might be expected or unexpected for a specific audience in a specific context (e.g. *see* the personalized privacy assistant based on privacy profiles in [187]), an approach which is also supported by the regulators. Such information can be then leveraged to create layered notices that are tailored to audience and context (*see* also [30]).

2.7.6 Unsolved Issues

Despite the promising novelties introduced by the GDPR's transparency obligation and active consent, there are issues that will remain unsolved. Firstly, the GDPR mandates an even longer list of information items to be presented in privacy notices than before. It reaffirms the paradigm according to which increased transparency is achieved by augmenting the quantity of information. On the one hand, privacy terms should be quick and easy to read. On the other hand, however, it is hard to convey all of the critical information for informed decisions in a coincise manner [197], while meaningful details risk to disappear. Indeed, some processing practices are so complex and the actors involved so many that some scholars [265, 211, 152] doubt that it will ever be possible to explain them in a simple and condensed notice, especially when users have little or no technical knowledge.

This is the so-called "transparency paradox" [211]: increased transparency is usually antithetical to simplified short notices. "Regarding the tension between comprehensiveness and comprehensibility, there *see*ms to be a blind spot between information requirements put in place for the benefit of consumers, and robust accountability mechanisms put in place in order to facilitate controls by supervisory authorities" [71, p. 58]. Indeed, more realistically, privacy notices are regulatory tools for other actors than the humble user: lawyers, regulators, journalists, advocates, investors, and industry [152].

Despite the criticism, examples showing that it is possible to provide complex information in small amounts of space exist (*see* e.g. [11] or [4]). The multi-layered notices proposed by the WP29 that will be analyzed more thoroughly in Section 5.2 can be very helpful to reconcile conciseness and comprehensiveness and to address different audiences. With the reaffirmation of transparency and consent, it is particularly important to enhance disclosures and consent mechanism because much burden is imposed on the individual [148]. However, adding icons and simplifying the terms cannot constitute the final and only solution to all the hurdles of the information paradigm explored in this chapter, although interested users and regulators can still benefit from them, as other research on visualizations shows. For example, one can hypothesize that there are some particularly sensitive contexts (e.g. online purchases or health-related services) where a user is more prone to consult the privacy statements.

Another issue that increased transparency cannot solve is linked to the modern complexity of data processing (*see* earlier Section 2.2.1.7). At the moment of data collection, hence at the time when privacy disclosures are provided to data subjects, the purposes for which data is collected are unknown or not defined yet. Other regulatory actions, for example the principle of data minimization, can counterweight the risks associated to this practice.

As the GDPR reaffirms consent as legal bases for processing, consent fatigue (*see* Section 2.2.2.1) will also not be solved. In fact, with the rising number of interconnected IoT devices collecting personal data, this situation is destined to worsen. Suffice it to say that in the very few days preceding the date of application of the GDPR, individuals received dozens and dozens of e-mails to re-consent and re-read the updated privacy policies of the services used: "[m]ass emailing, mass privacy changes, mass pop-ups on every websites [...] resulted in many individuals expressing their annoyance with having to accept and review the updated privacy policies, and the companies seemed to convey the message 'we are really sorry you have to go through this, but we are obliged by law to send you this spam' " [71, p. 10]. This evidence highlights that, despite the undebatable progress made, fatigue and information overload are not solved by the GDPR's transparency provisions.

2.8 Conclusive Remarks

This introductory chapter has, firstly, analyzed the origins and the reasons behind the information paradigm in data protection law. The provision of transparent information about data processing and data subjects' rights is deemed necessary to rebalance information and power asymmetries between the entity collecting personal data and the person providing that data. Moreover, disclosures should also help data subjects to take mindful decisions about their privacy, for instance whether to give consent to certain data processing or whether to choose one service over another.

However, much research shows that privacy policies are not usually read or are not understood, while consent is rarely informed, due to a variety of factors: the language is usually overly complex or vague, the length of the documents excessive, whilst the mostly non-existent structured layout discourages readership and comparability across notices. Other hurdles are represented by the omnipresence of notices and consent requests on websites, apps, and devices: the quantity is simply unmanageable for the single user. Also the fact that commonly timing of disclosures is decoupled from the timing of privacy-related decisions prevents data subjects from making informed choices.

Hence, there is a discrepancy between what the law presumes and how individuals behave in practice: although data subjects could access information about the processing of their personal data to inform their decisions, in practice they do not. This reality derives from the fact that the law treats human beings are rational decision-makers, whereas much empirical research has provided evidence about their cognitive shortcuts and biases. The way human reasoning works can be leveraged to predictably steer individuals towards privacy-preserving behaviors through choice architecture (i.e. nudges), for instance by setting privacy-friendly defaults or by highlighting risks related to personal data provision. Nevertheless, nudges can also influence data subjects in the opposite sense, for example to direct them to share more personal information. In Chapter 5, this topic will be discussed with practical examples and recommendations.

The GDPR seems to have integrated some behavioral insights to respond more realistically to users' bounded rationality and biases, and to guide them towards more privacy-conscious behaviors. Most prominently, the transparency obligation is not solely based on the quantity of information provision, but also dedicates an unprecedented attention to the quality of such information. Layout, presentation, visualization, and quality of language can in principle counteract the deficiencies of privacy-related communication identified by many researchers. User-centeredness is a key term: information and other experiences must be designed with the intended user in mind, and also rely on empirical research that provides evidence about the effectiveness of such approaches.

This framework profoundly informs the research presented in this dissertation, that focuses on the design of privacy-related communication and, specifically, on visual communication. The GDPR unprecedentedly acknowledges the potential of visual elements to simplify and clarify lengthy, cumbersome legal notices and suggests machine-readable icons to provide an overview of the intended data processing. The next chapter will provide a summary of the technologies that can support machine-interpretability of legal information, while the following chapters will focus on design-related issues.

Chapter 3

Technologies for the Representation of Legal Knowlegde

At the end of the last Chapter, the idea of machine-readable icons was introduced as a strategy suggested by the regulators to increase transparency of privacy information. It is from this provision, contained in Article 12 GDPR, that the research described in this and the following chapters originates. The GDPR's incentive to machine-readable, visual communication about data practices identifies two parallel and intertwined lines of research: the first direction is related to the technologies for the management and (semi-)automated extraction of legal information (explored in this chapter), whilst the second is linked to the user-centric design of legal information (illustrated in the next chapters). The final aim is the transformation of legal documents into both human- and machine-understandable formats.

In particular, the following sections detail how a "visual layer" can be automatically created from marked-up privacy policies, with the aim of communicating data practices in a human-friendly manner. Good practices drawn from information design, graphic design and legal design must inform the shaping of the visual privacy policy. As extensively documented in the next chapter, indeed, user-friendly legal documents based on visual and information design can, compared to traditional texts, enhance users' comprehension and address habituation effects. The opportunities offered by the digital environment have shifted the consideration of privacy policies as paper, text-only, static legal documents into interactive and user-centered interfaces to the legal content.

The research described in the following aims to bridge the gap between humans and computers: if ease of readability of the law can be pictured as a continuum, on the one edge are humans and on the other edge are computers, while text is equally distant from both endpoints. Whereas visualizations ease human accessibility, code ameliorates machines' accessibility to information [138]. Without an interface, machine-readable information is confined to the exclusive world of computers and technical experts, whereas user-friendly and visualized documents are not meaningful for machines. The final aim, as also advocated and envisioned by [138], is the transformation of legal documents into both human- and machine-understandable format: the existing fracture can be recomposed to generate legal information that is both. Moreover, the integration with automated technologies makes the hereby suggested approach scalable and applicable to big quantities of information, whereas most human-centered design applied to legal content is crafted *ad hoc* and will remain unique.

Standardization is also key: not only because familiarity with a certain standardized visual language lowers the chances of misinterpretation of graphical symbols [163], but also because there are considerable initial costs to generate images expressing different legal functions (as it will also be maintained in the last chapter). However, if this process is standardized, then it can even become automatized and the "initially higher coding costs will be reduced yet further, and accompanied by significantly lowered transmission, retrieval, and de-coding costs" [42, p. 44]. Several times during these pages, in this and the next chapters, the importance of scalable and replicable solutions will be highlighted.

Generally, Information and Communications Technologies (ICTs) can contribute to the purpose of creating more accessible and comprehensible legal documents for the human computation and, as a consequence, they can play a significant role to make the law more human-centered. For example, information technologies can contribute to make rules and remedies more accessible to citizens, therefore more effectively protecting their rights [259].

These opportunities flourish in the Semantic Web: machines process text according to its semantic content by enriching legal documents with machine-readable specifications that enable machine to machine interoperability: XML mark-up is used to embed meta-textual information into the legal document, complemented by languages, like OWL, that define conceptual structures and provide machines with the knowledge they need in order to understand the information they read (see Sections 3.1 and 3.2). There are multiple manners of representation for the same information: pure text can be transformed into a machine-readable representation, that can be leveraged, in its turn, to generate visual elements in a replicable manner (see Section 3.3). The concepts of a specific domain, alongside their corresponding visual representation, can be formally codified in an ontology, which is linked to the metadata mark-up of a legal document and that will be described in Section 3.4. Specific semantic content of privacy terms can be thus semi-automatically visualized with icons to make these documents more informative and human-centered.

3.1 Multi-Layered Legal Documents

Documents must be conceived as the actual textual content plus the additional information that describes and gives meaning to the document or to specific parts of the document. In the legal sphere, the online publication of documents has traditionally placed an emphasis on the layout and display of the information, with the attempt to replicate paper documents in a close manner. Although this is a fundamental aspect for the human eye to apprehend and make sense of information, it is not a meaningful element for machines that process the document because they do not own the same human knowledge or skills to interpret this information and draw inferences. On the contrary, computers must rely on an explicit, machine-interpretable (i.e. machine-readable) description of the structure and the meaning of a document and its parts. Such additional information is encoded by mark-up languages such as XML (eXtensible Mark-up Language). It is, thus, not a matter of displaying online documents that have been till hitherto in paper format, but rather a semantic-oriented presentation that allows machines to understand the content of such documents, with all the related benefits in terms of large-scale distribution, accessibility, and analysis.

Additionally, any document can be described through metadata, i.e. structured information that does not make part of the document, but that has been added to the actual content to enable machines to interpret it.This specific kind of structured data is described and organized according to an ontology, i.e. "an organized description of the metadata values that describe the resources" [287, p.41].

In addition, legal documents (i.e. legislation, regulations, contracts, etc.) are the sources of norms, guidelines, and rules that regulate behaviours and impose constraints on what is allowed or forbidden [34, 35]. If such knowledge is structured in a machine-readable format, operations of search, exchange, comparison, evaluation and reasoning become possible at a large scale. Such heterogeneous information contained in legal documents can be formally represented through a multilayered architecture. Indeed, different perspectives of analysis can be expressed through strictly separate layers [225]:

1. Text: the textual layer provides the representation of the document's original content;

2. Structure: the structural layer provides a hierarchical organization of the text,

3. Legal Metadata: the metadata layer connects document's information

with external ontological resources

4. Legal Ontology: the ontological layer is the formal model of the concepts mentioned in the document

5. Legal Rules: the logical layer provides the legal interpretation and modelling of the legal meaning of the text, and the transformation of the norms into legal rules to allow legal reasoning.

3.2 Standards for the Electronic Exchange of Information

It is good practice to rely on existing web standards, such as XML, URIs, XML schemas, RDF, OWL, etc., to represent and guarantee the validity of legal information over time [225] and to ensure interoperability. Standardization means crafting simple, technology-neutral representations of legal and legislative documents that enable the uniform structuring and the effective exchange of machine-readable information across different countries and jurisdictions [287]. An XML standard has the function of capturing and describing the existing similarities among documents, notwithstanding their differences [287]. Standards can, thus, enable open access on the generation, presentation, accessibility, and description of any document. A good standard does not only unify documents where they present similarities, but it also allows for individual differences to be expressed. This is why, flexibility and extensibility are essential features of standards. The extensibility that is needed to accomodate specific needs can be balanced through the rigid separation of the documents' information into the different layers introduced above. Such a layered partition safeguards the integrity of the original legal document over time, that can be managed without any modification to the authentic text [36].

The LegalDocML Technical Committee[1] adopted in 2017 Akoma Ntoso[2] (*see* Sect. 3.2.1) as legal open XML standard for legislative, judiciary and legal documents, that implements the first three levels of legal information described above. As a standard, Akoma Ntoso has been adopted to represent a varieties of documents by the European Parliament, the European Commission, the Parliament of Uruguay, the Italian Senate, the High Court of Cassation of Italy, the Kenya Law Report, the FAO, among the others[3].

The fourth level is implemented by the W3C Web Ontology Language (OWL) [6], which is a computational logic-based language that is able to represent rich and complex knowledge in documents known as ontologies and whose capacities will be discussed in Section 3.2.2. OWL is part of the W3C's Semantic Web technology stack. Lastly, the LegalRuleML Technical Committee[4] is promoting LegalRuleML [34] as an XML-based rule interchange language that can represent the legal rules pertaining to the fifth layer of legal information, described in Section 3.2.3.

3.2.1 Akoma Ntoso

The Akoma Ntoso schema is a technologically-neutral XML description of parliamentary, legislative and judiciary documents (e.g. legislation, debate records, etc.) that enables the addition of descriptive structure (i.e., the structural and semantic mark-up) to the content of such documents [225]. Akoma Ntoso implements the first three levels of the multi-layered architecture described earlier: it provides a vocabulary to capture structural and semantic elements of the legal document, but it also grants mechanisms for the reference to external ontologies and legal knowledge modelling [225]. At the first level, textual mark-up signals to the machine that a certain string

[1] https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legaldocml

[2] http://www.akomantoso.org

[3] For a complete list, visit http://www.akomantoso.org/?page_id=275.

[4] https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legalruleml

of character is, for instance, a date, a paragraph, or a heading. Moreover, it provides the vocabulary to express references to internal and external sources.

At the structural level (the second layer), the Akoma Ntoso vocabulary is based on patterns commonly found in legal documents and captures the semantic organization of the legal text: for instance, it identifies the preamble and the articles of a legislative act and provides these sections with the appropriate, machine-readable meaning. The analysis in the last Chapter pointed out that structure is a fundamental feature to ensure text navigability, while the next two chapters will outline how visual hierarchical can be used in legal documents.

3.2.1.1 Metadata in Akoma Ntoso

At the metadata level (the third layer), descriptions about the content of a legal document can be added [224]. This layer allows the association of the underlying levels with external ontological information (*see* next Section). Such machine-readable information is leveraged by semantic reasoners that are thereby enabled to extract meaning from the text and apply inference rules [36]. A part of the metadata is provided in a separate block (i.e. the metadata section), whereas another part is placed in the text (i.e. inline elements). Some other elements have no structural role, but are essential to provide strings or spans of text with semantic meaning. The tags of such semantic mark-up can be linked to a reference in the metadata section that points to an external resource defining the meaning of such tags [225], which allows the migration from the term to the concept [45]. In the legal sphere, semantic mark-up assume a particularly important role because the words (i.e. textual strings) appearing in a document must be unambiguously and clearly defined [37].

The metadata section provides information about the document's publication, its identification, its lifecycle, its presentation and other details [226]. A great relevance assumes the reference section: references are mechanisms that connect the document with external resources, thus connecting the third and the fourth level of the multilayered architecture (*see* Section 3.2.2).

Furthermore, the Top Level Classes or TLCs are highly generic groupings of instances, for which no particular meaning nor property is defined [36]. The lack of a strict definition for TLCs allows a certain degree of freedom in the choice about the ontology, that can be used to manage the concepts of the document. It is necessary, however, to bind the abstract TLCs to an external ontological model. Indeed, any legal document can be adapted to any ontological representation of concepts [36]. The network of meanings defined by the law and other legal sources can, thus, be changed according to the needs of the users and the applications, but an external resource ensures that the document is not consequently changed: only the vocabulary does.

Akoma Ntoso defines a minimal, loose ontology based on few TLCs: Person, Role, Concept, Organization, Object, Event, Place, Process, Term and Reference. It is up to the document's editors to associate a formal semantics to each class through a specific formalism (e.g. OWL). References can link the different linguistic realizations of concepts appearing in a document to the corresponding unambiguous instances of an external ontology [224]. For instance, the words "you", "data subject" and "user" appearing in a privacy policy all refer to the same ontological class of "data subject", which provides machine-interpretable semantic definition according to a specific legal framework. The semantic mark-up disambiguates identical textual expressions that, in fact, refer to different legal concepts. For instance, "consent" can be the action of giving consent, but also the legal document that gives lawfulness to the processing operations for which consent has been agreed.

3.2.2 Legal Ontologies

The ontological level (the fourth layer) is the semantic resource that aims to formally represent a domain of reality to enable the sharing of information and knowledge about it [45]. Legal ontologies are controlled vocabularies and can be expressed in many languages, like the ontology representation language (OWL). A classical definition from artificial intelligence characterizes the ontology as an explicit, formal specification of a shared conceptualization [126]. In other words, an ontology consists in an abstract model of concepts (i.e. a conceptualization), usually concerning a certain domain knowledge. The meanings of the concepts are defined (i.e. it is explicit) and transformed into a machine-interpretable format (i.e. it is formal). Moreover, a consensus about the modelling of the ontology has been reached (i.e. it is shared). Ontologies are massively adopted in the Semantic Web to support the uniform description, and the consequent retrieval and sharing, of legal knowledge, i.e. of legal concepts and their linguistic realization. They are organized in classes, i.e. abstract groups representing a certain concept: e.g. the class "data subject"; relations among classes, e.g. data subject is a subclass of the class role, and has rights, so it is linked to the class defining rights; and instances, i.e. individuals of a class.

Depending on their domain coverage and on their goal, there exists several typologies of ontologies. Foundational ontologies bear great relevance for any domain, since they aim to remove terminological and conceptual ambiguities, while they define domain-independent top level classes aimed to be reused in the design of any other domain-specific ontology. These categories can be considered basic construction blocks that define standardised knowledge representations to guarantee semantic interoperability: foundational ontologies define highly general concepts and relations among concepts that can be identically described across any domain, e.g. the concept of event. Without such a shared, basic knowledge, different applications would interpret the same concept differently, resulting in a computational Babel tower that reaches the opposite result to the desired one: the impossibility of knowledge sharing. There also exist core ontologies that define the key elements of the vocabulary of a specific domain and are used as basis for the conceptual specialization in the domain ontologies. Hence, a network of semantic resources to describe a certain area can be created: a domain-specific ontology, like the ontology for data protection described below in Section 3.4, is based on one or more legal core ontologies that define basic legal entities (e.g. legal roles or events), which is in turn linked to a foundational ontology. Moreover there are also ontology design patterns, that are small, motivated ontologies used as modeling components in ontology design that fulfil the goal of reusability [119]. All of these types of ontologies have been employed to build PrOnto (*see* Section 3.4).

3.2.3 Legal Rules

The fifth level of the multilayered architecture is represented by legal rules, whose logical structures can be formally modelled through the LegalRuleML XML-based rule interchange language [35]. This language can capture the distinctive features of legal and legislative documents and integrates the other layers, thus completing the levels of representation of the Semantic Web. Indeed, basic XML is not able to express key elements of a legal resource, such as to operate a classification of norms, by differentiating constitutive from prescriptive rules. The former provide definitions of a certain concept, e.g. " 'personal data' means any information relating to an identified or identifiable natural person" (Art 4.1 GDPR). The latter regulate actions or the outcome of actions, e.g. "the controller shall be able to demonstrate that the data subject has consented to processing of his or her personal data" (Art. 7.1 GDPR). Indeed, deontic norms regulate behaviours and impose constraints on what is permitted or forbidden: obligations, permissions, prohibitions, rights. Not only: inside a norm, different parts with different values can be distinguished, such as bearer and conditions [35].

Thus, legal texts' provisions expressed in natural language can be converted into non-ambiguous logical representations of their meaning. This is possible only if there exists a logical representation of the prescriptive and constitutive rules of a certain domain knowledge that is aligned to the items and relations of a specific ontology. Only by operating such conversion into a machine-readable, logical format, compliance with respect to a certain input can be automatically assessed. For instance, the controller has the obligation to comply with GDPR's requirements about lawfulness described in Article

6, which establishes that data processing must have a specific purpose and be based on one of the possible legal bases. An automated reasoner can, thus, analyze privacy policies and consent agreements to determine whether a consent request constitutes a valid legal ground for processing in a specific context. A logical modeling of the legal documents can even capture the consequences for the data subject if he does not provide his data in specific cases, as stated in Article 13, e.g. "BlaBlaCar will not be able to provide you with the services offered on our Platforms if the required information is not provided, consequently you will not be able to register for a user account on our Platforms[5]".

3.3 From Law to Code to Machine-Readable Visualizations

Although the above described semantically-enriched representation allows for data exchange and automated reasoning, there is yet another layer that can be added on top of the composite legal information architecture: a visual layer. Indeed, the machine-readable representation of the document content allows the semi-automated display of graphical elements that can potentially increase individuals' understanding of data practices. Machine-readable privacy policies and consent forms can be regarded in this sense as interactive digital interfaces between controllers and data subjects.

In the last few years, there have been many applications of information visualization that have improved how computers present data to people in a human-readable format, but not enough integration of such activity with knowledge visualization. Knowledge visualization is mostly related to knowledge transfer among humans. There can be an integration between the two: ameliorating how humans communicate knowledge to other humans through a computer-mediated language and through visualizations.

[5] Extracted from the BlaBlaCar's privacy notice at https://www.blablacar.co.uk/about-us/privacy-policy.

3.3.1 Formalization of Data Protection Information

Following the GDPR guidance on the use of icons for transparent communication (*see* Chapter 1), a data protection icon set has been designed and will be properly described in Chapter 6. The design of the icon set has followed a methodology that combines legal design and formalization of legal knowledge [223]. In this approach, the first step consists in the identification of the information items that are intended to be rendered visually: since it is semantic information (and not, say, logical or numerical) the appropriate representation is pictorial. Unlike other data protection icon sets (*see* Section 6.1.4), such conceptual circumscription derives from an analysis of the legal requirements that need to be respected for regulatory compliance. The GDPR explicitly details how organizations must fulfil the obligation of mandated disclosure. Article 12 defines the condition for transparent communication and grants permission to use icons, rather than other types of visual elements. Articles 13-14 enumerates the exact pieces of information that must be provided in privacy notices, e.g. the purposes and the legal basis for the processing. In this light, the law provides exact and clear indications, whereas a human-centered design approach would favor an initial analysis of users' needs to steer the design process and would encourage experimentation with multiple communicative techniques (*see* Chapter 4).

From the analysis of legal requirements derives the design of a computational ontology. This shared formalization lowers the chances of personal biases' influence on the representation of legal knowledge and its interpretation. The entities formally represented in this semantic resource can be linked to their corresponding icons, thus creating an interconnected network of machine-readable and visual representations (*see* Fig. 3.1). Such binds can provide the visual elements with a precise, stable, and machine-interpretable meaning as the GDPR explicitly requires. It can be argued that enriching the icons with such extra information can open up possibilities to develop tools that make use of such extra information, and that can presumably limit the chances of misrepresentations and misinterpretations of the graphi-

Legal reference from Art. 13	**Concepts**
1(a)	Controller
1(c)	Purposes of processing
	Legal bases for the processing
1(e)	Recipients
1(f)	Transfer to third country
2(a)	Storage period
2(b)	Right to access Right to rectification Right to erasure Right to restriction of processing Right to object to processing Right to data portability
2(c)	Right to withdraw consent
2(d)	Right to lodge a complaint with a supervisory authority
2(f)	Automated decision-making

Table 3.1: Taxonomy of the mandatory information extracted from Art. 13 of the GDPR. The first column lists the article's points, the second details the corresponding concepts.

cal symbols, while also allowing the provision of the same content in multiple languages.

Machine-readability even allows for the semi-automatic retrieval of the visualizations encoded in the ontology, once that the semantic expressions in natural language of the text have been associated to their corresponding ontological representations through the XML mark-up. As a consequence, the icons can be semi-automatically retrieved to appear in correspondence of the matching privacy terms contained in the document, summoned by the Akoma Ntoso semantic tags, provided that a dedicated application is designed and implemented. In the hypothesis underlying this research, derived from the studies on legal visualizations that will be introduced in Chapter 4, this would make privacy statements on specific topics easier to find and understand. Icons can also visually represent those pieces of information that shall appear when consent is asked, such as the processing operations (e.g. transfer, anonymize, erase, etc.) and purposes (e.g. marketing, profiling, etc.) (*see* e.g. Figg. 3.6 and 3.7).

Finally, this formalized conversion of natural language into a conceptualization and, in turn, into visualizations, provides traceable passages of the transformation from text to machine-readable to visual. Not only: if the visualizations are machine-readable, they can be conceived and designed as any other kind of legal information, without regard to their nature. For instance, the metadata section of documents that have been marked up through Akoma Ntoso and LegalRuleML embed in the document important information on the annotation (for instance, about the authoritativeness of the annotator) [225]. Metadata can thus also refer to the images contained in a legal document and be leveraged in the presentation. Exact references on the intended meaning and use of the image can be directly embedded in the image itself.

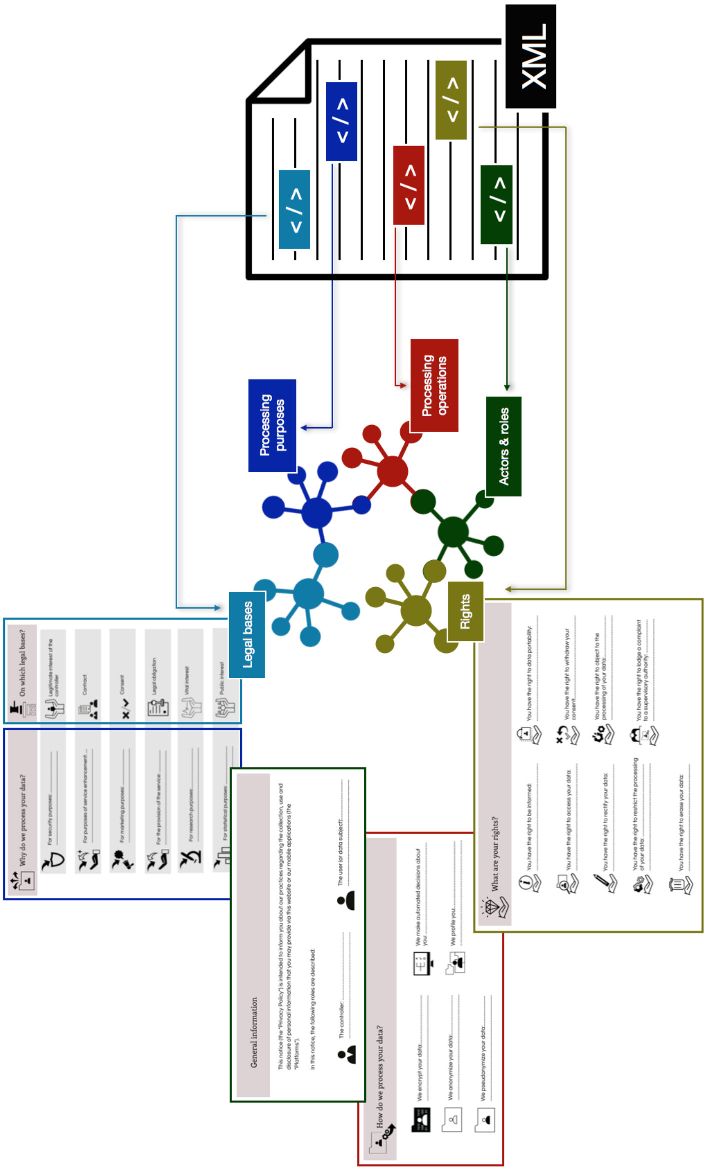


Figure 3.1: The semantic network that exists between the concepts in the ontology, the corresponding Akoma Ntoso mark-up of a privacy policy, and a visual layer with information architecture and icons.

3.3.2 Related Work on Semi-Automatic Legal Visualizations

The legal visualizations that will be described in Chapter 4 are mostly *ad hoc*, unique creations. However, the approach presented here attempts to deliver repeatable and scalable solutions - in this light, in the next two chapters, patterns will be discussed at length. There have been some experiments around the automatic visualization of contract terms (e.g. payment clauses or contract duration clauses)[235]. The visualization is achieved through an online graphical interface[6] that allows users to select among multiple choices or to fill in fields with the necessary information (e.g. the start and end dates of the agreement). This tool generates a few different kinds of design patterns depending on the data typology, such as timelines or graphs, and is intended to offer a ready-made solution to help legal drafters to make the meaning of their contracts clearer.

The digital ecosystem and the semantically-enriched legal information offer additional opportunities: encoding data in a machine-readable format allows the semi-automatic generation of visualizations. For example, XSLT (XML Style Language for Transformations) [66], a W3C recommendation, is used to transform XML structured documents (e.g. Akoma Ntoso document) into other documents (e.g. HTML documents) that will be interpreted and displayed by the presentation engine (e.g. a browser) [277]. The same input data can thus be rendered differently, depending from the device and application.

Some public institutions recognize the value of visualizations to represent legal information, namely to offer a human-interpretable interface to make sense of that data that, otherwise, would simply look like complex and copious machine-interpretable data. In Italy, for example, the legal open data set released by the Regione Piemonte and marked-up in Akoma Ntoso was used

[6] https://cs.anu.edu.au/people/Michael.Curtotti/visualcontracting/#close. Accessed on June 21, 2018.

to display the complexity of legal order overtime[7] [219]. Also marked-up legislative datasets from the Italian Chamber of Deputies and from the Italian Senate were visualized[8] with the purpose of, for instance, the comparison of legislative activity among different legislatures, by displaying the complexity of the legislative procedure and by visually quantifying the time that is necessary for the adotpion of bills. Such applications represent perfect examples of how the integration between legal visualization and legal informatics can facilitate the management of the complexity of legal knowledge [120].

3.3.2.1 Semi-Automatic Visualizations of Privacy Information

As for what concerns the privacy and data protection domain, only a few researches have been conducted so far to build semi-automatic visualizations on machine-readable data. For instance, an interface based on icons showing how well a situation matches users' privacy preferences was developed on the P3P machine-readable format [72], whereas an extension of the ODRL permission-based language[158] was proposed to support visualization through icons across different social network providers [159].

Recently, an attempt to machine-readable representation of privacy notices has been made in the US by the Usable Privacy Policy Project[9] (UPP) [258], that has similar goals to the project described in this dissertation. The project aims to semi-automatically extract key concepts from privacy notices in order to allow both user-tailored presentations and large-scale analysis of privacy policies [296]. The marked-up privacy policies are used to train Natural Language Processing models that learn to automatically extract relevant information from privacy policies and to constitute the semantic foundation of the PrivOnto ontology [217], which formally represents the data practices described in the privacy policies. This approach is different from PrOnto (*see* Section 3.4) mainly because UPP consists in a bottom-up approach: it derives the annotation scheme from common elements found across a pool

[7] http://lodpiemonte.cirsfid.unibo.it. Accessed on June 21, 2018.
[8] http://code4italy.cirsfid.unibo.it. Accessed on June 21, 2018.
[9] https://explore.usableprivacy.org. Accessed June 21, 2018.

of privacy policies, which however vary greatly because in the US there is no regulatory framework at the federal level. Besides, the privacy policies of the UPP corpus are addressed to American users. This is why, the taxonomy lacks a number of terms that are proper of the EU legal framework. The GDPR, conversely, defines in Articles 13-14 the items of information that must be mandatory disclosed to data subjects, regardless of the fact that they are actually included in an organizations' privacy notice. The codification of such information in PrOnto paves the way to a top-down approach that can, in principle, allow the comparison of privacy statements expressed in natural language with an ideal, general model of a GDPR-compliant privacy policy.

Recently, an approach based on deep learning analysis of privacy polices was proposed. Polisis (i.e. Policies analysis) [145] is a framework that automatically annotates a privacy policies at a very fine-grained scale according to the schema of labels defined in [297] and explained earlier (*see* also Section 5.2). A combination with icons is attained to evaluate the accuracy of Polisis in the automatic annotation of fragments of a privacy notice, namely to assign an icon expressing a specific meaning to the corresponding span of text. Polisis makes use of the TRUSTe icons [281], which assume different colors (i.e. green, yellow, red), depending on the level of permissiveness of a certain data practice. In this sense, this attempt is remarkable, because it reaches high levels of accuracy and can not only signal if a certain practice is present in the privacy notice, but also a description of its fairness. However, such categories do not cover the extent and the purposes of the GDPR, and the display of the icons might be controversial because it is based on a completely automated analysis of the legal text, combined with a judgment on the fairness of the terms.

A similar approach, but with a European focus, is represented by Claudette, a machine learning and natural language processing system that can detect unfair privacy terms under the GDPR, in terms of comprehensiveness, clarity and lawfulness [71]. A similar approach could be leveraged to semi-automatically annotate privacy policies with PrOnto's concepts because it can automatically detect certain concepts expressed in natural language, such as legal bases, processing purposes, etc., and annotate them with the relevant tags. This automated semantic annotation could then be used to display the corresponding graphical symbol as our approach proposes. However, at the state of the art, these methods currently do not achieve impressive results in precision and recall, indicating that final human intervention is needed to approve or disapprove the automatic mark-up.

If we exclude the abandonded P3P [288], at the present day, to the best of our knowledge, there is no standard that is specifically intended to represent the information that is proper of privacy policies in a machine-understandable manner. Since more and more personal information is shared among interconnected devices (IoT) and this tendency is destined to grow, machine-readable privacy policies may gain acceptance [261, 25]. Research towards the combination of machine-readable policies and pictographic approaches has been encouraged [155]. Despite all the good reasons mentioned, the fear is that any sort of technological mediation (like XML encoding or pictorial representation) between privacy policies and users can result in inaccurate or biased information, on which nevertheless users will base their decisions. Unlike traditional textual privacy policies, that are legally enforceable because they convey the same message to every user [59], machine-readable legal documents risk introducing ambiguity and legal uncertainty. However, solutions like public comprehensive documentation, standard certification, and guidelines on the interpretation of machine-readable formats have been proposed [182]. Moreover, automated visualization is not meant to replace the original privacy policy, but it rather "decouples the legally binding functionality of privacy policies from their informational utility"[145, p.2].

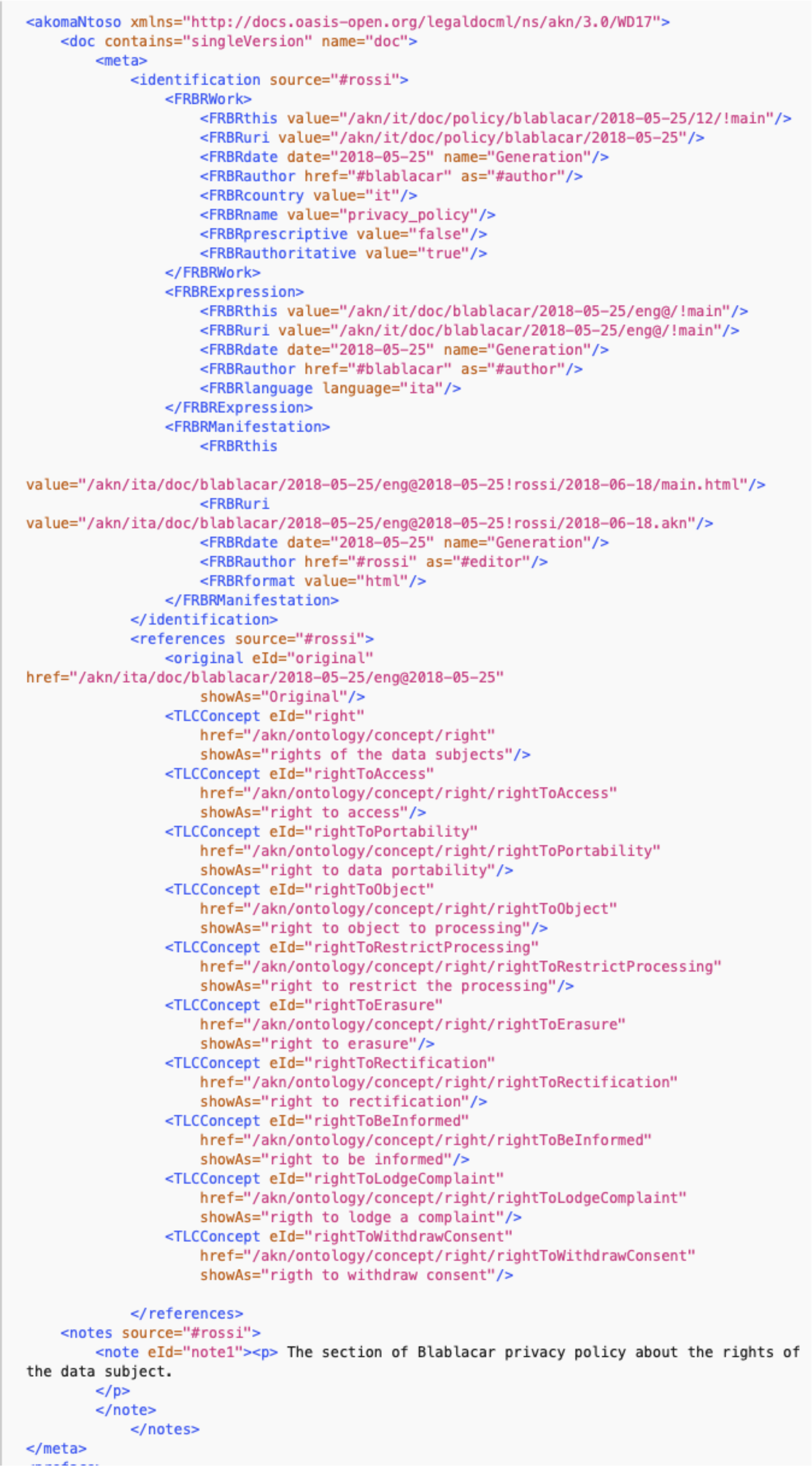
3.3.3 Modelling Privacy Policies

For all of these reasons, the research described in these pages has relied on the Akoma Ntoso standard, that was used to mark up the case study for the present research: the BlaBlaCar's privacy policy[10] that was judged complete with respect of the innovations introduced by Article 13, at the time of annotation. For instance, the processing purposes are present in quantity and are clearly associated to a legal basis. Also the rights of the data subjects are thoroughly explained. Thus, this privacy policy was deemed suitable for the visualization because, unlike others, offered many concepts that have corresponding icons in the set, and are well-organized in the relevant sections.

An annotation schema derived by the ontological formalization was developed to mark-up the information about the data practices. The document class that provides the highest degree of flexibility was selected, because it provides an open structure for those types of documents that are poorly standardized and whose structure is highly varied [226], unlike e.g. acts or judgments. Indeed, although there are regulatory requirements about the content and the language, the structure of privacy policies is not fixed. The document's organization, such as headings, paragraphs, etc., is made explicit through standard elements. Finally, the instances of the ontological concepts in the privacy policy were mapped to the corresponding TLCs in the metadata section.

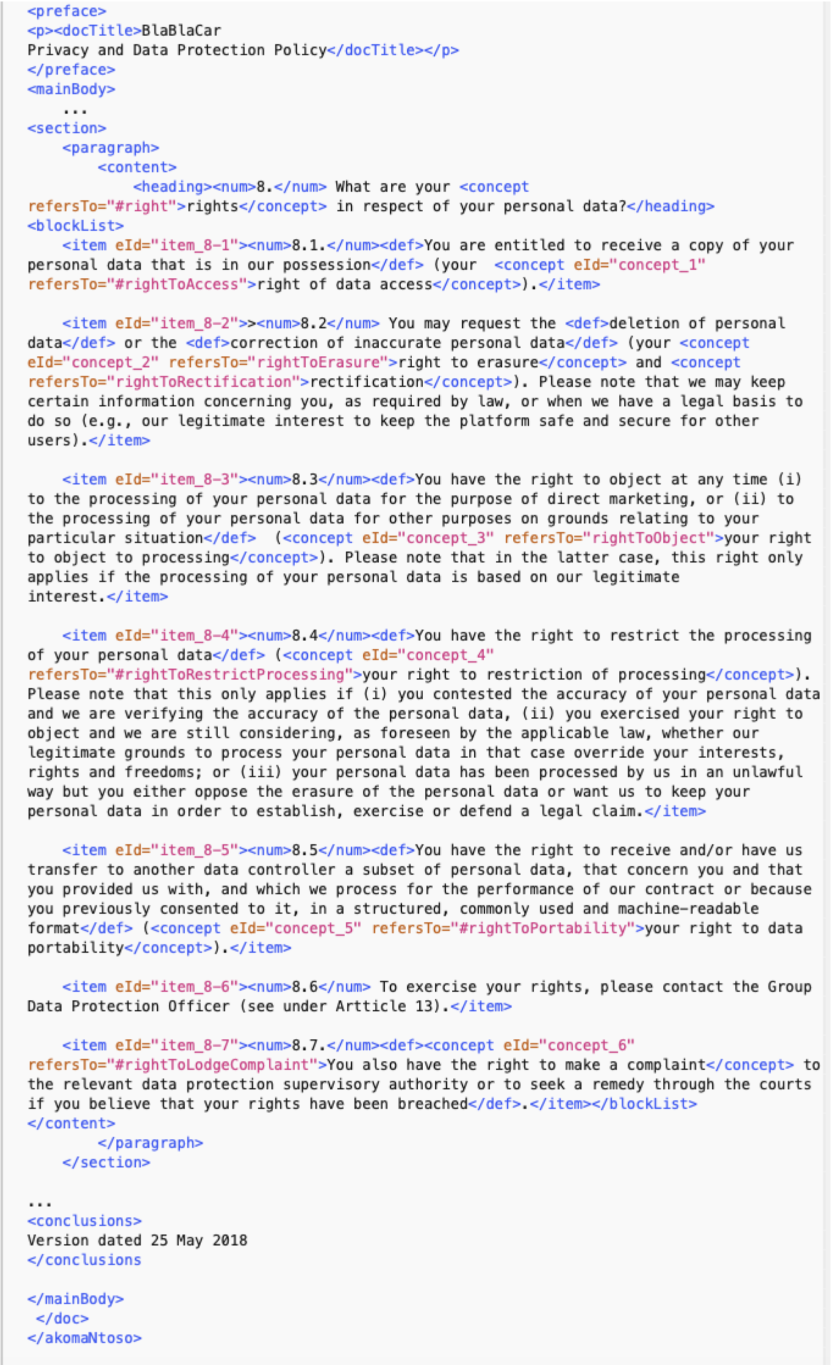
In the following XML excerpt (Fig. 3.2 and Fig. 3.3), an illustrative example of the privacy policy's section about the rights of the data subject is provided. The concepts (i.e. the rights) expressed in natural language are marked up with a concept element, whose attribute "refersTo" points to the corresponding reference contained in the metadata section. The IRIs of the TLCs have informative content because the Akoma Ntoso naming convention has the purpose of creating URIs that univocally identify the concepts, but also provide some information about the concepts without accessing the external ontology, e.g. about parent-child relationships (e.g. right as parent class of the subclasses of the different rights).

[10] https://www.blablacar.co.uk/about-us/privacy-policy



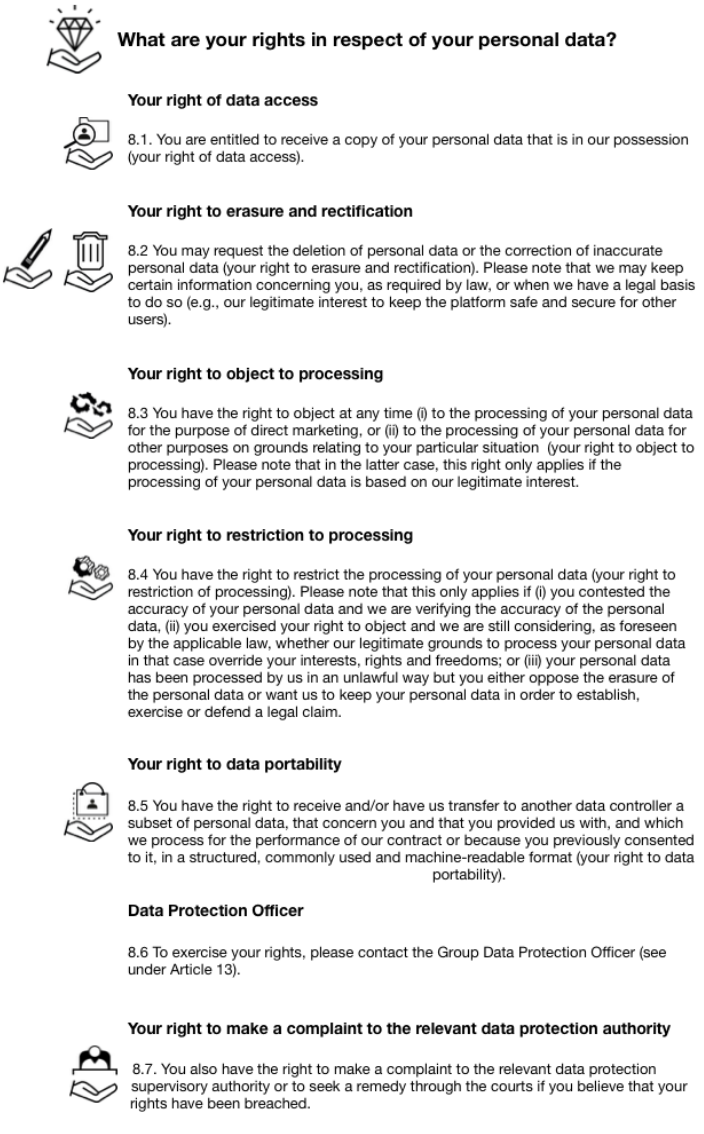
```
<akomaNtoso xmlns="http://docs.oasis-open.org/legaldocml/ns/akn/3.0/WD17">
    <doc contains="singleVersion" name="doc">
        <meta>
            <identification source="#rossi">
                <FRBRWork>
                    <FRBRthis value="/akn/it/doc/policy/blablacar/2018-05-25/12/!main"/>
                    <FRBRuri value="/akn/it/doc/policy/blablacar/2018-05-25"/>
                    <FRBRdate date="2018-05-25" name="Generation"/>
                    <FRBRauthor href="#blablacar" as="#author"/>
                    <FRBRcountry value="it"/>
                    <FRBRname value="privacy_policy"/>
                    <FRBRprescriptive value="false"/>
                    <FRBRauthoritative value="true"/>
                </FRBRWork>
                <FRBRExpression>
                    <FRBRthis value="/akn/it/doc/blablacar/2018-05-25/eng@/!main"/>
                    <FRBRuri value="/akn/it/doc/blablacar/2018-05-25/eng@/!main"/>
                    <FRBRdate date="2018-05-25" name="Generation"/>
                    <FRBRauthor href="#blablacar" as="#author"/>
                    <FRBRlanguage language="ita"/>
                </FRBRExpression>
                <FRBRManifestation>
                    <FRBRthis

value="/akn/ita/doc/blablacar/2018-05-25/eng@2018-05-25!rossi/2018-06-18/main.html"/>
                    <FRBRuri
value="/akn/ita/doc/blablacar/2018-05-25/eng@2018-05-25!rossi/2018-06-18.akn"/>
                    <FRBRdate date="2018-05-25" name="Generation"/>
                    <FRBRauthor href="#rossi" as="#editor"/>
                    <FRBRformat value="html"/>
                </FRBRManifestation>
            </identification>
            <references source="#rossi">
                <original eId="original"
href="/akn/ita/doc/blablacar/2018-05-25/eng@2018-05-25"
                    showAs="Original"/>
                <TLCConcept eId="right"
                    href="/akn/ontology/concept/right"
                    showAs="rights of the data subjects"/>
                <TLCConcept eId="rightToAccess"
                    href="/akn/ontology/concept/right/rightToAccess"
                    showAs="right to access"/>
                <TLCConcept eId="rightToPortability"
                    href="/akn/ontology/concept/right/rightToPortability"
                    showAs="right to data portability"/>
                <TLCConcept eId="rightToObject"
                    href="/akn/ontology/concept/right/rightToObject"
                    showAs="right to object to processing"/>
                <TLCConcept eId="rightToRestrictProcessing"
                    href="/akn/ontology/concept/right/rightToRestrictProcessing"
                    showAs="right to restrict the processing"/>
                <TLCConcept eId="rightToErasure"
                    href="/akn/ontology/concept/right/rightToErasure"
                    showAs="right to erasure"/>
                <TLCConcept eId="rightToRectification"
                    href="/akn/ontology/concept/right/rightToRectification"
                    showAs="right to rectification"/>
                <TLCConcept eId="rightToBeInformed"
                    href="/akn/ontology/concept/right/rightToBeInformed"
                    showAs="right to be informed"/>
                <TLCConcept eId="rightToLodgeComplaint"
                    href="/akn/ontology/concept/right/rightToLodgeComplaint"
                    showAs="rigth to lodge a complaint"/>
                <TLCConcept eId="rightToWithdrawConsent"
                    href="/akn/ontology/concept/right/rightToWithdrawConsent"
                    showAs="rigth to withdraw consent"/>

            </references>
    <notes source="#rossi">
        <note eId="note1"><p> The section of Blablacar privacy policy about the rights of
the data subject.
        </p>
        </note>
            </notes>
</meta>
```

Figure 3.2: The metadata block of the XML mark-up on the section about data subjects' rights in the Blablacar's privacy policy



```
<preface>
<p><docTitle>BlaBlaCar
Privacy and Data Protection Policy</docTitle></p>
</preface>
<mainBody>
    ...
<section>
    <paragraph>
        <content>
            <heading><num>8.</num> What are your <concept
refersTo="#right">rights</concept> in respect of your personal data?</heading>
<blockList>
    <item eId="item_8-1"><num>8.1.</num><def>You are entitled to receive a copy of your
personal data that is in our possession</def> (your  <concept eId="concept_1"
refersTo="#rightToAccess">right of data access</concept>).</item>

    <item eId="item_8-2">><num>8.2</num> You may request the <def>deletion of personal
data</def> or the <def>correction of inaccurate personal data</def> (your <concept
eId="concept_2" refersTo="rightToErasure">right to erasure</concept> and <concept
refersTo="rightToRectification">rectification</concept>). Please note that we may keep
certain information concerning you, as required by law, or when we have a legal basis to
do so (e.g., our legitimate interest to keep the platform safe and secure for other
users).</item>

    <item eId="item_8-3"><num>8.3</num><def>You have the right to object at any time (i)
to the processing of your personal data for the purpose of direct marketing, or (ii) to
the processing of your personal data for other purposes on grounds relating to your
particular situation</def>  (<concept eId="concept_3" refersTo="rightToObject">your right
to object to processing</concept>). Please note that in the latter case, this right only
applies if the processing of your personal data is based on our legitimate
interest.</item>

    <item eId="item_8-4"><num>8.4</num><def>You have the right to restrict the processing
of your personal data</def> (<concept eId="concept_4"
refersTo="#rightToRestrictProcessing">your right to restriction of processing</concept>).
Please note that this only applies if (i) you contested the accuracy of your personal data
and we are verifying the accuracy of the personal data, (ii) you exercised your right to
object and we are still considering, as foreseen by the applicable law, whether our
legitimate grounds to process your personal data in that case override your interests,
rights and freedoms; or (iii) your personal data has been processed by us in an unlawful
way but you either oppose the erasure of the personal data or want us to keep your
personal data in order to establish, exercise or defend a legal claim.</item>

    <item eId="item_8-5"><num>8.5</num><def>You have the right to receive and/or have us
transfer to another data controller a subset of personal data, that concern you and that
you provided us with, and which we process for the performance of our contract or because
you previously consented to it, in a structured, commonly used and machine-readable
format</def> (<concept eId="concept_5" refersTo="#rightToPortability">your right to data
portability</concept>).</item>

    <item eId="item_8-6"><num>8.6</num> To exercise your rights, please contact the Group
Data Protection Officer (see under Arttticle 13).</item>

    <item eId="item_8-7"><num>8.7.</num><def><concept eId="concept_6"
refersTo="#rightToLodgeComplaint">You also have the right to make a complaint</concept> to
the relevant data protection supervisory authority or to seek a remedy through the courts
if you believe that your rights have been breached</def>.</item></blockList>
</content>
        </paragraph>
    </section>

...
<conclusions>
Version dated 25 May 2018
</conclusions

</mainBody>
 </doc>
</akomaNtoso>
```

Figure 3.3: The XML mark-up of the section about data subjects' rights in the Blablacar's privacy policy



What are your rights in respect of your personal data?

Your right of data access

8.1. You are entitled to receive a copy of your personal data that is in our possession (your right of data access).

Your right to erasure and rectification

8.2 You may request the deletion of personal data or the correction of inaccurate personal data (your right to erasure and rectification). Please note that we may keep certain information concerning you, as required by law, or when we have a legal basis to do so (e.g., our legitimate interest to keep the platform safe and secure for other users).

Your right to object to processing

8.3 You have the right to object at any time (i) to the processing of your personal data for the purpose of direct marketing, or (ii) to the processing of your personal data for other purposes on grounds relating to your particular situation (your right to object to processing). Please note that in the latter case, this right only applies if the processing of your personal data is based on our legitimate interest.

Your right to restriction to processing

8.4 You have the right to restrict the processing of your personal data (your right to restriction of processing). Please note that this only applies if (i) you contested the accuracy of your personal data and we are verifying the accuracy of the personal data, (ii) you exercised your right to object and we are still considering, as foreseen by the applicable law, whether our legitimate grounds to process your personal data in that case override your interests, rights and freedoms; or (iii) your personal data has been processed by us in an unlawful way but you either oppose the erasure of the personal data or want us to keep your personal data in order to establish, exercise or defend a legal claim.

Your right to data portability

8.5 You have the right to receive and/or have us transfer to another data controller a subset of personal data, that concern you and that you provided us with, and which we process for the performance of our contract or because you previously consented to it, in a structured, commonly used and machine-readable format (your right to data portability).

Data Protection Officer

8.6 To exercise your rights, please contact the Group Data Protection Officer (see under Article 13).

Your right to make a complaint to the relevant data protection authority

8.7. You also have the right to make a complaint to the relevant data protection supervisory authority or to seek a remedy through the courts if you believe that your rights have been breached.

Figure 3.4: A possible 'visual layer' for the Blablacar privacy policy's section about data subjects' rights, with icons and information architecture

3.3.4 A Privacy Visual Layer

Once that the structural elements and the data protection concepts of the privacy policy have been marked up, visualization modalities must be designed in order to achieve the purpose of delivering the company's data practices more clearly and more engagingly through the help of visual elements. The inspiration for this research originates in document and information design's best practices and successful visualization experiments. Below, we introduce visual aids to the comprehension and navigation of the privacy policy, that can stem from the legal XML mark-up explained earlier.

3.3.4.1 Layout and Structure

In the first place, XML allows for structured, semantically-enriched layout. Document layout is crucial to strengthen users' desire to read privacy policies (see Chapter 2). Document layout improvements include dividing a long text into small chunks of information (e.g. the different policy's sections) in order to make it more digestible and fight users' discouragement. On small device screens, information can be split instead of displayed in a unique scrolling window. Moreover, typography can be enhanced, by using user-friendly fonts and, particularly, by using colors, font size, and bold typography to signal hierarchy of information and to bring attention to the different topics of the privacy policy. These elements can highlight a path through the document, which enhances its skimmability and it allows easier and quicker retrieval of specific pieces of information. Information hierarchy can be conveyed also by changing font size and thickness of sections' and subsections' headings. Furthermore, hyperlinks or pop-up windows can be added to critical legal notions of the privacy policy (and even to the icons, *see* below), that might be unclear to laypeople (e.g. controller, data processor, recipients, etc.). Pop-up windows or links to explanations and examples can be provided in order to support the comprehension. In this way, it is possible to create a network of information that is not made exclusively of text, but also of visual elements of various kinds.

3.3.4.2 Icons

Other kinds of visualizations can be implemented to display relevant semantic information concerning data practices (*see* Fig. 3.4). Standardized icons that refer to key principles of data collection and processing can act as information markers and thereby enhance strategic reading, as it will be discussed at length in the next chapters. They can be reused in every privacy policy to enhance comparability among them. Simplified comics and flowcharts (potentially also animated visuals) can be used to represent data interactions between the stakeholders: controllers, processors, data subjects, etc.

3.3.4.3 Legal Rules

Finally, appropriate ways of representing legal rules must be envisioned: since LegalRuleML can represent the logical structure of norms, a visualization that leverage on the different deontic norms (i.e. obligations, permissions, rights and prohibitions) and on their components (i.e. bearer, conditions) can be designed. In the first place, logical implications can be formalized and consequently visualized to make them more relevant. For instance, Article 13 highlights the importance of apprising data subjects about the consequences of retaining their personal data (*see* Fig. 3.5). Similar types of visualizations can leverage the logical layer to clearly show the conditions under which certain rights or obligations apply and can make use of the icons, as in Fig. 3.6. In addition, the consequences of giving or withholding consent can be visually represented and logically linked to certain conditions ("if...then..."). Let's assume the frequent case where, if the data subject gives his consent, then she will receive marketing communication. One icon depicting consent could be linked via a consequential relation to an icon depicting marketing communications, as displayed in Fig. 3.7. Over time, individuals get accustomed to the visualizations, especially if standardized, and use them as quick and effective ways to find information.

Visualizing the differences between what users have the right to do or shall

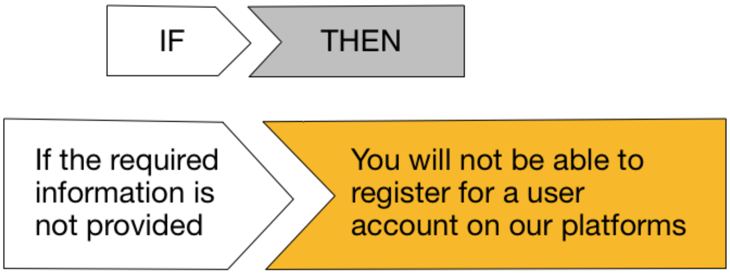

IF
THEN
If the required information is not provided
You will not be able to register for a user account on our platforms

Figure 3.5: A possible visualization of a logical implication that highlights what happens if the data subject refuses to provide her data, extracted from the BlaBlaCar's privacy policy.

do under the privacy policy terms might reasonably give them a straightforward and quick way to understand at first glance whether the company's data practice correspond to their preferences. If, for instance, a service obliges the user to hand over many typologies of data, and some of them also unexpectedly, and this characteristic of the data practices is highlighted by a simple up-front visualization, it can be easier for users to decide whether or not to accept the data practices. On the contrary, if the visualization highlights users' rights, users might reasonably believe that the company has an ethical approach towards their data. Controllers' obligations and rights could be put *vis-a-vis* with data subject's rights and obligations in a swimlane to ease comparison (*see* Section 4.6), whereas conditionals and consequences of specific choices can be represented as diagrams and flowcharts (*see* Figures 4.7 and 4.8). Besides, regulators could leverage on visualizations to easily and quickly understand which policies are legally compliant on a large scale.

3.4 PrOnto: the Privacy Ontology

As anticipated earlier, semantic tags acquire meaning only if an ontological representation is linked to them. This is why an ontology of the regulation on data protection concepts and relationships is currently under development: PrOnto, the Privacy Ontology [222, 221, 220]. This ontology

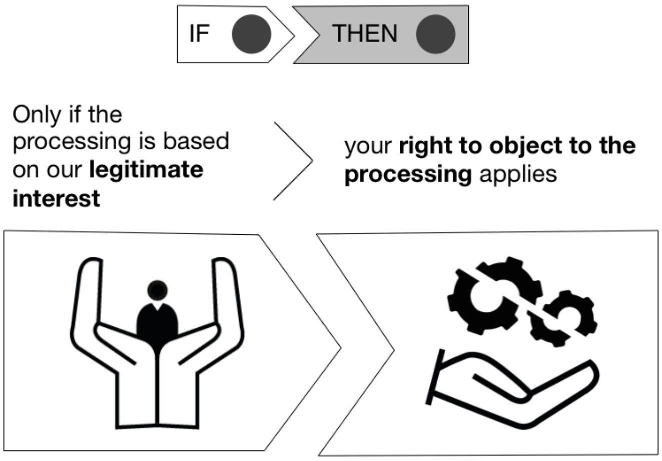

IF
THEN
Only if the processing is based on our **legitimate interest**
your **right to object to the processing** applies

Figure 3.6: A possible visualization with icons describing the conditions under which the data subject has the right to object to processing, extracted from the BlaBlaCar's privacy policy.

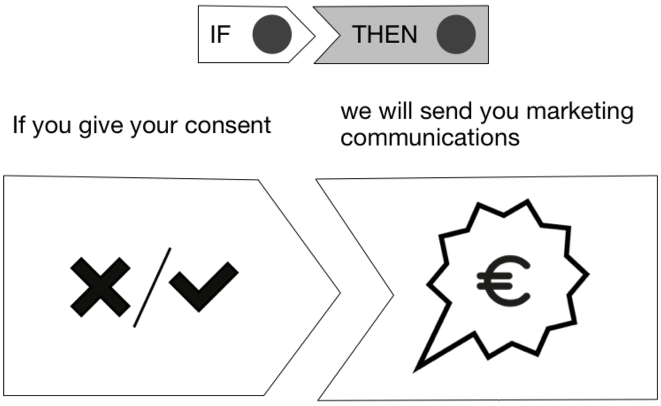


Figure 3.7: A possible visualization of a consent form, with the icons depicting the consequences of giving consent.

is mainly focused on the GDPR, because, unlike other ontologies focusing on privacy and data protection, it is mainly aimed at helping companies and organizations to comply with the provisions set forth by this specific Regulation. The explicit goal of this data protection ontology, combined with other Semantic Web technologies and legal reasoners, is to ease the data controllers' fulfillment of duties such as the adherence to privacy-by-design principle, the undertaking of the data protection impact assessment and the detection of those violations (e.g. a data breach) that need countermeasures. However, the GDPR constitutes only the initial, central core of norms, that can be expanded to other legal frameworks and jurisdictions.

3.4.1 The Design of PrOnto

PrOnto has been designed by following MeLOn [222], an interdisciplinary Methodology to build Legal Ontologies, which is composed of a series of recursive steps.

In the first place, the specific goals of the ontology are defined, i.e. the research questions that the ontology aims to address and the practical use-cases where the ontology might be helpfully applied. In PrOnto's case, this means modeling the legal norms defined in the GDPR to allow legal reasoning and compliance checking. Thus, PrOnto has put an emphasis on the modeling of the processing operations and of the obligations and rights belonging to the different actors (e.g., data subject, controller, etc.) defined by the Regulation.

The normative text, i.e. the GDPR, was analyzed by the team's legal knowledge engineers to extract relevant concepts and relations among them, e.g. the different stakeholders affected by this Regulation and their respective rights and duties. This knowledge was then integrated with expert feedback and additional information taken from other authoritative sources, such as Opinions and Guidelines from the Article 29 Working Party (e.g. [27]) and guidance from the UK's Information Commissioner's Office (e.g. [161]), as well as international standards (e.g. [169]).

Moreover, the best practices for ontological knowledge modeling introduced in Section 3.2.2 have been followed, therefore PrOnto is framed in foundational and core ontologies such as ALLOT [38], FRBR [5], LKIF-core [50], and PWO [118]. For the functionalities of PrOnto, also ontology design patterns expressing values in time and context [238] have been reused.

MeLOn also establishes the evaluation of the ontology applied to concrete use-cases in terms of coherence, completeness, efficiency, effectiveness, agreement, and usability. Lastly, a testing phase that makes use of the OntoClean method [129] and of SPARQL queries establishes if the research goals defined at the beginning of the ontology design have been reached. Publication and feedback collection is the last step that contribute to reach a shared agreement within the community of legal experts.

3.4.2 PrOnto's Conceptual Modules

The Privacy Ontology is a composition of conceptual modules organized around fundamental data protection aspects that the GDPR governs:

1. data and documents;

2. agents and roles;

3. data processing and workflow;

4. processing purposes and legal bases;

5. legal rules and deontic operators.

For the scope of the data protection icon set that will be described in Chapter 6, only the relevant modules will be introduced here, whereas the general ontological architecture and its single modules have been discussed at length in [222, 221, 220]. The single classes will not be defined in this chapter, but will be rather described when the icons will be introduced.

3.4.2.1 Data and Documents

The focus of the Regulation is personal data, i.e. it is the object of its protection, but it is also the origin of the relations among different actors, such as the data subject, the controller, the processor, the supervisory authority, and so on. Such relations can be described and regulated through documents: privacy policies, contracts and the consent subclass, codes of conduct, laws and case-law. For these reasons, the FRBR ontology design pattern adopted for the publication process [5] has been applied to the modeling of data and documents. The GDPR, however, not only defines its scope, i.e. personal data, but also what falls outside of its scope, i.e. non-personal data and anonymised data. Moreover, the Regulation establishes different rights and obligations when processing special categories of personal data, i.e. sensitive data. In addition, personal data can be classified with respect to their origins and, initially, three broad categories of personal data were identified accordingly, following [13]: the personal data that the data subject provides directly or that is observed from her behavior; the personal data that has undergone some kind of processing; and the personal data that is inferred or derived and that is generated by the controller through the analysis of other data. This distinction was initially made because it was deemed crucial for the exercise of certain rights, as will be outlined below, although it has been abandoned in the current version of the ontology and replaced by a more complete and thorough modeling of the processing operations with a focus on the result that they produce (e.g. the pseudonymize operation produces pseudonymized data).

3.4.2.2 Agents and Roles

Agents and the roles they assume represent an additional fundamental conceptual core of data protection law. Such a distinction is central, but often overlooked, in legal ontologies: an agent may play multiple roles depending on the context and the processing operation. For instance, a person has the rights of the data subject when her personal data is processed, but has the obligations of a controller when collecting personal data of other people, and can even act as processor or third party in relation to other data processing. As it is evident, the adoption of different roles by the same agent triggers the exercise of different rights and duties. The first version of the icon set (*see* Section 6.2.2.1) comprised the roles of data subject, controller, processor, third party and supervisory authority.

3.4.2.3 Data Processing Operations

Evidently, the diverse typologies of operations that can be taken on personal data constitute a cornerstone of the data protection domain and its rules. The processing activities are modelled through a workflow [118], i.e. a sequence of steps with a specific input and a specific output, characteristic that is shared with many other human activities. The workflow execution is composed of actions, namely events that are specified by temporal and contextual parameters, such as place and jurisdiction. The essential actions that were identified (*see* Figure 3.8), and subsequently rendered graphically, are: anonymize (subclass of delete), pseudonymize (subclass of derive), automated decision-making (individual of infer), profiling, direct marketing, encrypt, copy, and transfer of personal data to third countries (individual of the class transmit, specified with a place axiom). Automated decision-making, profiling, and transfer to third countries, moreover, assume particular relevance because they must be prominently signaled and explained in any communication directed to data subjects, as Article 13-14 prescribe.

3.4.2.4 Processing Purposes and Legal Bases

The principle of lawfulness (Art. 6) establishes that personal data processing must be motivated by specific purposes, that were identified and extracted from the normative text (*see* Fig. 3.9): security purposes, research purposes, statistical purposes, profiling purposes, marketing purposes, judicial purposes, health-related purposes, humanitarian purposes, journalistic purposes, and purposes of public interest. Every purpose must be supported

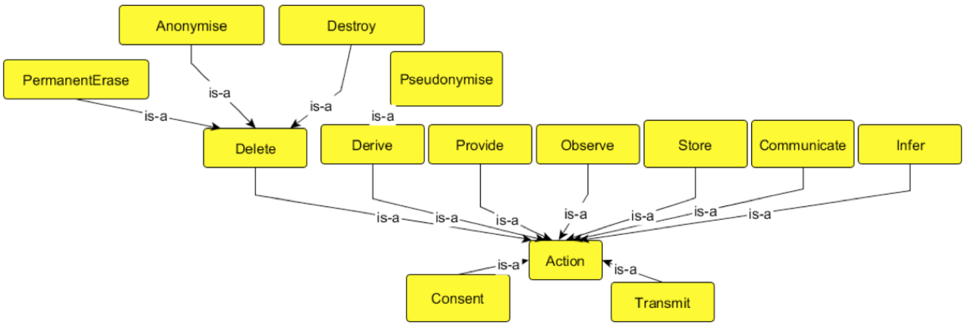


Figure 3.8: The PrOnto module on processing operations.

by one of the possible legal bases laid down in Article 6: consent, contract, legal obligation, public interest, vital interest, legitimate interest. Note that the consent and the contract are subclasses of the document class.

3.4.2.5 Legal Rules and Deontic Operators

Since one of the main goals of PrOnto is to support compliance checking with the GDPR, the modeling of legal norms in terms of deontic operators, i.e. rights, obligations, permissions, and prohibitions, bears considerable relevance and can be integrated with LegalRuleML. In the perspective of transparency, the rights of the data subject assume paramount importance (Articles 12-22). Therefore, the following subclasses are included in the ontology (*see* Figure 3.10): right to be informed, right to access, right to rectification, right to erasure (or "right to be forgotten"), right to data portability, right to withdraw consent, right to restriction of processing, right to object to processing, and right to lodge a complaint to a supervisory authority.

3.4.3 Icon Ontologies

Recently, a possible manner of formalizing the syntax and semantics of an iconic language for the medical domain with an icon ontology was proposed in [181]. The explicit goal of such formalization is to define icons' meanings for the creation and interpretation of new elements and to allow automatic

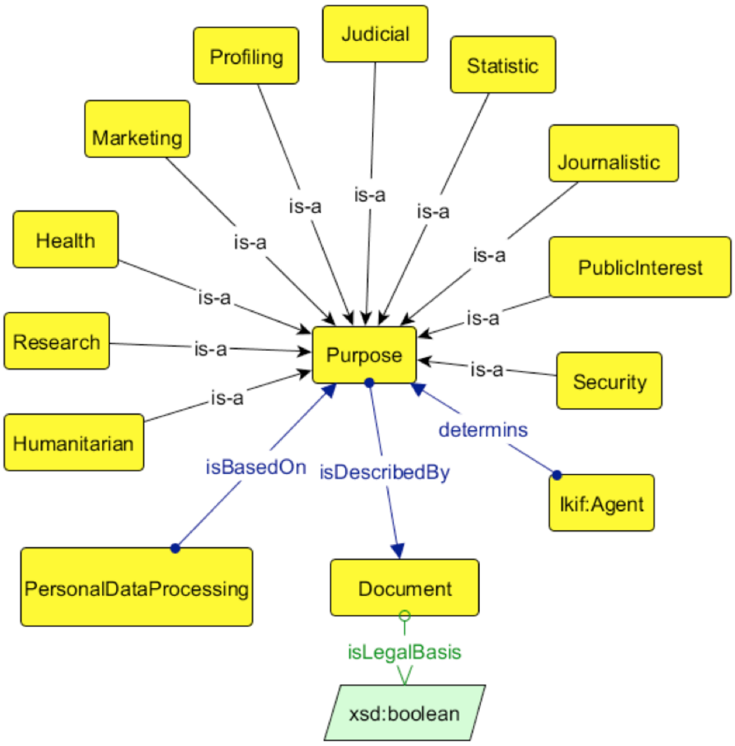


Figure 3.9: The PrOnto module on processing purposes.

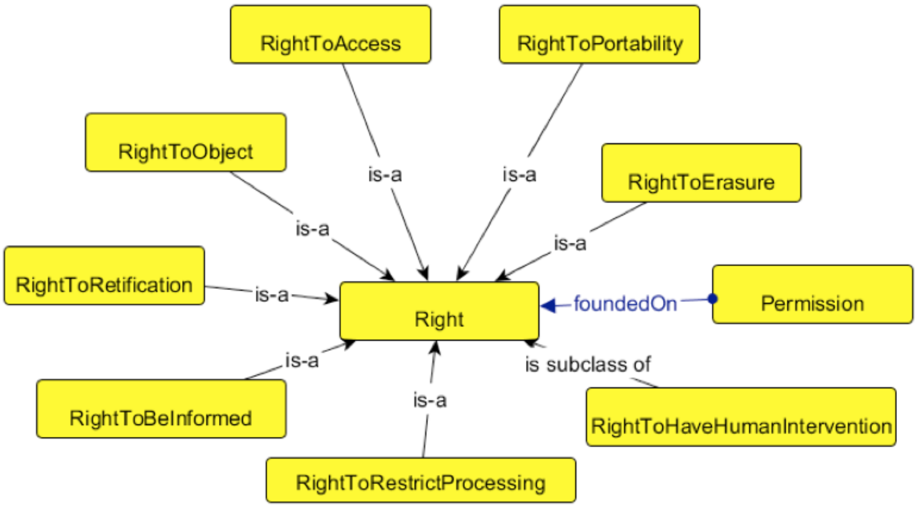


Figure 3.10: The PrOnto module on data subjects' rights.

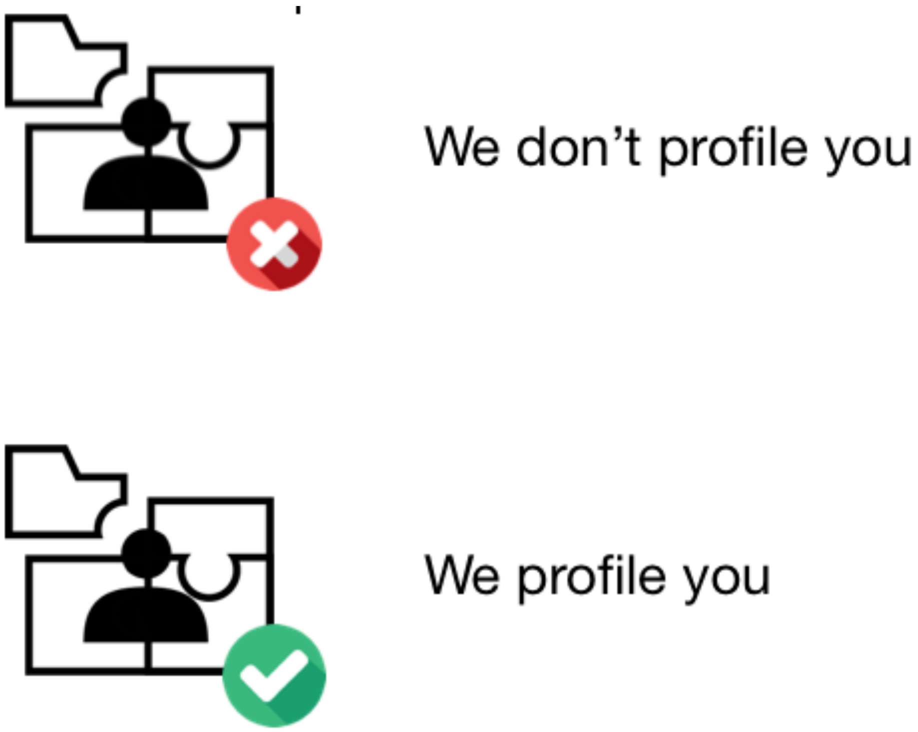


Figure 3.11: A possible visualization of the profiling icon combined with symbols indicating whether the data controller does or does not perform the practice.

tasks, e.g. the generation of multilingual labels, for the VCM iconic language. VCM combines shapes, pictograms, and colors to represent medical concepts (e.g. a patient's clinical conditions, her medical history, her treatments) and is meant to help health professionals to access medical documents in a more efficient manner. Thus, this approach can inspire future work for an effective mapping between an icon ontology and a data protection domain ontology, but it would result exceptionally useful for the coherent creation of new icons based on automated processing, for instance to represent whether a certain organization allows or does not allow a certain practice, e.g. profiling.

3.5 Conclusive Remarks

This chapter has described technologies commonly employed for the management of legal information. XML and ontologies can provide legal documents with machine-interpretable meaning that allows, for example, for automated reasoning. In the present approach, semantically-enriched privacy policies can be leveraged to generate a user-friendly visual layer composed of icons and structured layout that can ease the navigation, comprehension and comparability of these documents, as will be motivated in the next chapter. The integration of human-centered document design with semantic technologies can in principle make this approach scalable and repeatable, unlike other solutions that will be proposed in the next chapter, but that would greatly profit from standardization.

Chapter 4

Legal Design and Legal Visualizations

The research described in these pages is set in the domain of legal design, an emerging discipline that proposes a “design-driven approach to legal innovation” [141]. It is necessary to stress the fact that, given the novelty of this approach, academic bibliographical sources are still limited, although rapidly increasing. The scarcity of academic references for this knowledge area reflects its innovative nature and highlights the need for a definition framework that can account for what legal design is and what it is not.

Together with other authors we have therefore drafted a Legal Design Manifesto[1] that defines what a legal design approach to the law is. The main points will be argued in the next pages.

This is why, it is somehow easier to define the outcomes, results, and products of legal design, with the description of its several, multi-faceted concrete applications, than to find sources providing theoretical foundations to the discipline. However, the following Chapter will attempt to present both the theoretical perspective and the concrete outcomes, and will put an emphasis on the studies that have more profoundly and significantly influ-

[1] A first version of the Manifesto can be found on https://www.legaldesignalliance.org/ and is currently open to comments. Last accessed: October 30, 2018.

enced the present research.

Firstly, Section 4.1 will introduce legal and document literacy, in order to understand the features that would support individuals' sense-making of legal information. Secondly, in Section 4.2, the emerging discipline of legal design will be presented, with a focus on its rationale and its methods, together with the completely new user-centered perspective that it launches in the legal sphere. Thirdly, in Section 4.3, many examples of a fundamental tool of legal design will be described: legal visualizations. Section 4.4 draws a unifying line among law, semiotics, and design, to identify a communicative perspective that is common to all. Then, in Section 4.5, the possible difficulties that can arise from the interpretation of legal visualizations and the associated risk of misinterpretation will be discussed, since they inform the design of the data protection icon set that will be described in Chapter 6. Lastly, Section 4.6 will describe legal design patterns to enhance legal communication, i.e. re-usable forms of a solution to a commonly occurring problem. Such repeatable solutions assume particular relevance within the view of standardization, as legal technologies described in the last chapter and the icon set described in Chapter 6 attempt to do.

4.1 Legal and Document Literacy

Although some voices are extremely critical about the possibility of educating individuals to increase their awareness towards legal [215] and data protection matters [40] (see also Chapter 2), there are some important considerations to be taken into account. Firstly, as literacy empowers individuals to become full members of a written language community and, therefore, enables them to influence the reality around them, similarly legal literacy empowers people to become fully functional members of the society and enables them to know how to take appropriate decisions in law-related contexts [303].

Legal literacy, among the many definitions (*see* [303], Chapt. 1) can be described as “[t]he ability to understand words used in a legal context, to draw conclusions from them, and then to use those conclusions to take action” [33, p. 23]. This definition reveals the double nature of literacy: not only it concerns the understanding of legal information and legal issues, but it also includes the ability to act upon them [303]. Indeed, legal information must not be simply decoded, but rather it must be useful for those that are expected to take decisions based on it (e.g. give or withhold informed consent) or to act accordingly (e.g. the exercise of a right). Hence, it is important to find appropriate means and formats to make such information readily graspable, but it is also necessary to empower people to find that information, and to learn how to use it.

As digital literacy has been included in the basic skills of the European citizen, then it could be auspicable that every member of the society had at least basic literacy about the legal context she lives in. However, it can be argued that such an objective is too broad and probably unattainable: if literacy has failed, as the data reported in Chapter 2 suggests, legal literacy will even be more difficult to attain. Nevertheless, this reality should not indicate that any tentative is lost: new ways to inform and educate citizens are arising and the pervasive use of technological devices makes the law more tangible than ever before. Data gathering is pervasive in every life domain and, due to its very nature, it offers additional, new opportunities to inform and raise the awareness of data subjects through technological means. Finally, although the goal of legal literacy is challenging, it does not mean that it is unattainable. Specific means can be designed to accompany the sense making and the decisions of the data subject, for example visual elements.

For what concerns those documents that have the goal to transmit knowledge from a party that owns it (e.g. the data controller) to a party that does not have the information (e.g. the data subject), another type of literacy acquires importance: “ ‘document literacy’ - a form of literacy that goes beyond reading the words to include strategic reading - searching documents for answers to questions, assembling information from different documents, and determining the relevance of information" [291, p.9]. As such, documents are not considered as mere containers of information, but rather as functional tools that human beings use to achieve their goals. Indeed "[e]ffective documents are structured around users' strategic needs to access different information at different times for particular purposes" [291, p.9]. On the contrary, poorly designed legal documents have been realized without a user [291] and a function in mind, as it was maintained in Section 2.2.1 for what concerns privacy policies.

These features mean that the individual should be able to use privacy policies to get sufficient, navigable, relevant information about the consequences of providing her personal data for processing. Not only: such information must be timely to allow for action to follow and appropriate in quantity to avoid overburdening the individual. Data subjects should also be enabled to understand how to give and withdraw consent. In other words, the information provided in privacy notices should be not only easily graspable, but also actionable.

Even further: design itself should easily guide the data subject towards the exercise of her rights. For instance, when the GDPR came into effect, websites offering their contents and services to users on the European soil changed their way to ask for cookie consent. Some solutions proposed a simple and usable adjustment of cookies, where the user is given relevant information about the purpose of each typology of cookies and about the necessity or, conversely the facultativity, of giving consent, and its implications. If the provision of such information is combined with a toggle bar or a similar interactive technology, it allows users to easily apply that recently acquired knowledge. In other cases, on the contrary, instructions to disable cookies are decoupled from the possibility to do so, which makes this action very hard or even impossible. Other solutions repropose less actionable and less informative strategies: not a concrete possibility to opt out, but rather a link to a long page listing the cookies used on a website, and a link from every single cookie on the list to other pages where it is either possible to opt out or merely to learn about how to opt out. This example is important because it shows in practice the difference between providing information and enabling individuals to use such information to reach a specific goal. Specific strategies to erode data subjects' rights and actions are unfortunately in use and will be analyzed in Section 5.3.

4.1.1 Proactive Law

The rationale behind legal design and legal visualizations (*see* Sections 4.2 and 4.3) is rooted in Proactive Law, a movement that focuses on achieving positive goals and outcomes. It stresses the importance of the needs and relationships of all those who use the law, not only legal experts but also laypeople (*see* [43] and [236] for the relationship between proactive law and legal design) [42]. The goal of using visualizations in legal communication is to make the latter more cognitively accessible and functional - in this light, legal documents such as contracts become generators of value instead of weapons to use against one other in case of conflict.

Traditionally, contracts and terms of service are written for the lawyers that want to protect their clients in case of a legal dispute, rather than for the people that want the relationship to succeed [138]. Similarly, clear and transparent privacy policies should mirror the transparency of the data processing they describe to generate trust in the data subjects. However, the reality is that most privacy policies, similarly to contractual terms, "favor legal certainty and formal enforceability [...] over efficiency-enhancing communication" [42, p.25]: privacy policies can be conceived as empowering instruments that inform data subjects of their rights but also their duties (e.g. in terms of security), instead of tools that hide companies' responsibilities and discharge their legal liability.

As such, proactive law is more focused on the future than on the past: it rather tries to prevent problems and disputes from arising, than to resolve conflicts afterwords (i.e. it is preventive) and attempts to achieve desirable outcomes and benefits from the beginning, instead of dealing with the consequences of failure (it is proactive). Rules are not conceived only in terms of compliance and the focus is not only on minimization of risk, but there the goal is the achievement of successful relations and positive effects between the parties [130]. Such proactive attitude reveals many points in common with the privacy by design approach (*see* Chapter 5) and represents one of the main points of the Legal Design Manifesto mentioned earlier.

4.2 Goals and Methods of Legal Design

Legal design is an interdisciplinary field, at the intersection among law, design, technology and behavioral studies. It has been defined as “the application of human-centered design to the world of law, to make legal systems and services more human-centered, usable, and satisfying” [141, Chap. 1]. Human-centered design focuses on the development of solutions that consider the target audience’s needs and requirements and that aim to enhance the effectiveness and efficiency of an artefact (e.g. a legal document) or an experience, by ameliorating human well-being and user satisfaction [166]. Thus, the perspective of the people that will use a specific artefact to accomplish a certain task or goal becomes a central part of its development [14], this is why the users are involved at every stage of the design, from the conceptual phase to the evaluation - what is called participatory design. The underlying assumption is that the experience of the law, its world of legal rules and services can draw lessons, tools, and mindsets from design and be thereby improved. “[Legal design] offers intentionality in the face of a system that has been hacked and patched together haphazardly and without user testing” [141]. For this reason, legal design seems to recompose the fracture that was highlighted in Section 2.2.1: there is a considerable distance between theoretical assumptions of the law and concrete individual’s behaviors.

4.2.1 Users of the Law

Human-centeredness means considering the end-users of a legal artefact. In this light, two main stakeholders are involved: firstly, the legal experts that want to better practice the law. And secondly, but of utmost importance, the lay people that can be empowered to better understand how the rules apply to them and how to be more in control of the complexities of legal matters. Very often, this second key stakeholder is not traditionally taken into consideration when legal solutions are crafted: for instance, legal documents such as contracts and privacy policies are drafted with a legal professional as target audience, thus shaping the modes and means of communication on the model of legalese. However, such legal documents do not serve the needs of the other individuals that will need to read and understand them to consequently take decisions or initiate actions.

4.2.2 Plain Language

Given this perspective, it is easy to identify the first seeds of the legal design movement, at least for what concerns legal communication, in the plain language movement. Although the attempt to make the law more accessible is not exclusively modern, it is the publication of *The Language of the Law* by Mellinkoff [200] that heralds the beginning of this movement, followed by the creation of experts' associations (e.g. Clarity[2] and Plain Language Network[3]), handbooks for legislators, and campaigns for the simplification. As recalled in the first Chapter, the plain language has even entered the legislation and stringent legal requirements that explicitly ban obscure information addressed to laypeople are now present in national and Union law, such as in the GDPR, in the Consumer Rights Directive [227], and in the Unfair Terms Directive [216]. Nevertheless, the plain language approach is still subject to diehard conservatism and entrenched in traditional practices and habits, thus legal practitioners struggle to adopt it.

[2] http://www.clarity-international.net/

[3] http://plainlanguagenetwork.org/

The plain language movement revolutionizes the assumptions behind professional writing by proposing that any (legal, administrative, medical, academic, etc.) communication can be packaged in a clear, straightforward, and simple style, with the needs of the reader in mind [32], while remaining precise and technically sound. It means that the drafters should put themselves in the shoes of the intended reader, conceive what and how the reader will possibly understand what they write and rewrite anything that is potentially unclear, ambiguous or difficult. The plain language movement has the goal of overcoming the barriers of aloofness and obscurity, and the feeling of intimidation that traditional legal writing creates in non-experienced readers [56]. It is not simply a matter of more simple lexical choices: plain language interventions consider information organization (i.e. information architecture) and restructure the design and layout of the document as fundamental means to support the understanding of the reader.

4.2.3 Information Design

The plain language movement also shares principles with information design, an interdisciplinary discipline that combines together different research subfields of linguistics, psychology and design generally concerned with making information accessible and usable to people [264]. Information design puts an emphasis on the satisfaction of the information needs of the intended receiver of a message and therefore combines analysis, planning, presentation, and understanding of a given message, i.e. its content, language and form [240]. In face-to-face communication, the instant feedback that a speaker receives from her counterpart (e.g. another speaker or an audience), such as facial expressions or verbal utterances, allows her to adjust her discourse to realize the smoothest communicative exchange possible. On the contrary, the asynchronous written communication is characterized by a “feedback gap” [290] between writer and reader. It is therefore of utmost importance to focus attention and resources on the way the message can be more effectively packaged to achieve clarity and to ensure that its intended receiver can make sense of it.

4.2.4 Areas of Intervention

Information and visual design constitute the core of legal design in those cases where it focuses on human-centered communication, probably the most developed subarea (*see* the numerous examples in Section 4.3.4). Yet, legal design encompasses other branches of design, according to the problem it tries to solve: the design of tools to accomplish a certain task related to work more efficiently (i.e. product design), the design of better experiences for individuals that face and go through legal processes (i.e. service design), the design of legal practices that make legal professionals work better and accomplish better outcomes (i.e. organization design), and finally the overarching design of complex systems that better serve the people and deliver value (i.e. system design) [141, 231]. These areas of intervention taken together represent a composite ecosystem that employ different tools and methods, and involve people with different skills, to achieve specific goals. For such reasons, the scope of legal design is defined accordingly: it can be a document (e.g. a contract), a product (e.g. an app), an experience (e.g. a trial), an organization (e.g. a corporation) and even a complex system (e.g. a new court).

Despite the variety of its applications, legal design in general is an approach to the assessment and creation of legal services that focuses on usability, utility, and engagement [141]. In ISO's words [166], usability is defined as the "effectiveness, efficiency, and satisfaction with which specified users achieve specified goals in particular environments". It is distinct from the mere functionality of an object, which is what the object can do [231], because it adds the human component and the interaction. For instance, a printer can be able to print documents (i.e. its functionality), but it can be unnecessary complex to employ for the user (i.e. its usability). Effectiveness depends on whether the user is able to complete a certain task, efficiency concerns the amount of effort to complete a certain task, whereas satisfaction indicates whether the user's feelings are positive (e.g. gratification) or negative (e.g. frustration) during the interaction with an artefact [19].

4.2.5 User-Centeredness and Participatory Design Methods

As it becomes evident, user-centeredness is key: the intended users of a certain artefact should be involved from the early design phases to provide an understanding of the problem that considers the multi-faceted views of different stakeholders. This is of extreme importance especially when professionals, such as lawyers, should devise a solution: the curse of knowledge and experience can profoundly impact the ways experts shape solutions by replicating what they already know or take for granted. Indeed, research has established that people tend to project their own beliefs and assumptions on others [42]. Hence, legal innovators should be able to separate these two levels, i.e. the own and the other, and approach any problem with open-mindness and criticism - for such reasons, interpersonal collaborations are encouraged and even considered necessary. On the one hand, this allows the understanding of the users, i.e. the target of the design, that become part of the design team as "experts on their own experiences" [286, p.10]. On the other hand, collaboration seeks to overcome personal bias and to reach mutual understanding among stakeholders that might have very different backgrounds and expectations.

These collaborations usually take place in workshops, and more specifically in (legal) design jams[4], where a concrete goal (i.e. a real-world problem) is provided to the participants who, in a limited time span, brainstorm solutions and embody them in prototypes that will be afterwards refined by professional designers [112]. All this is typical of participatory design, where users are involved at every stage of the design process and become co-designers [14]. Chapter 6 will provide more indications about participa-

[4] http://legaldesignjam.com/

tory design methods, by providing practical examples and setting them in the specific context of this research, i.e. the creation and evaluation of a data protection icon set.

4.2.6 The Design Process

Contrary to common ideas, design is more about the process of creation and innovation than the outcome of such process, i.e. it puts an emphasis on the conception of new ideas and not on the finished product [42]. It is for this reason that legal design welcomes the methodology of problem-solving based on design thinking and migrates the lessons learned in the business sector to the realm of law: design thinking is a human-centered approach to innovation [51] that focuses on the process to developing practical, creative solutions to problems. Like legal design, design thinking is based on user-centeredness, multidisciplinary collaboration, visual communication, and prototyping [231]. "[D]esign thinking leads to the generation of new ideas, but also validates them through analysis and evaluation" [42, p. 14]: indeed, innovation and user research constitute core aspects of legal design.

4.2.6.1 Analysis

Solutions arise when the actual behavior of individuals and their problematic interaction with artefacts are directly or indirectly observed. This is why stakeholders should be carefully identified: individuals have different cognitive and information needs, depending on their expertise, background knowledge, level of education, culture, age, etc. For instance, it can be safely assumed that lawyers and legal authorities will use a privacy policy to achieve different goals than ordinary website's visitors: the former to monitor the lawfulness of an entity's data practices, the latter to look for a few specific information items that interest them. Therefore, a concrete analysis must be carefully carried out to discover specific, as opposed to general and abstract, needs and opportunities, for instance by directly observing how individuals interact with a certain artefact, similarly to ethnographic research, or by conducting e.g. background interviews and questionnaires. This is called generative research and has the goal of exploring the context to inspire and inform the design team in the early stages of the design process [286].

Indeed, user research is a collaborative exploration between designer and user [42] and can make explicit individuals' needs, expectations, and desires [14]. Its function is the identification of opportunities where interventions for the enhancement of the *status quo* are needed. In the research described in this dissertation, the analysis stems from a literature review since the typical issues of privacy policies have been already well documented. User research can more precisely define the weak points of a certain legal artefact that, therefore, offer opportunities for improvement. Such a mapping of the *status quo* guides the subsequent design stages: the definition of the problem and of the potential areas of intervention originates the generation of ideas and solutions.

4.2.6.2 Synthesis

The phase of observation and understanding paves the way to the phase of invention. It is important to stress that there is no unique, winning idea to solve a specific problem: openness and experimentation are principles that every designer should keep in mind. Legal design does not aspires to become a prescriptive theory that generates "a single 'right' procedure, image or layout" [42, p.22], but rather a creative and iterative process that can provide indications about the chances of suitable use of a certain element, e.g. a visualization, given information type and goal of the design. Unlike other contexts, there is no predefined and known objective: objectives might even change as the understanding of the problem evolves with the solution [42] and as users' needs arise, sheding light on overseen aspects. Also, it is good practice to examine the existing landscape, to borrow and re-use promising ideas (e.g. design patterns, *see* Section 4.6) and to build on the inspiration they provoke or on their evident shortcomings [141].

This synthetic phase of the design process is characterized by experimentation and iteration [42]. Interpersonal engagement is key at this stage to broaden one's own exclusive perspectives and to create a number of prototypes of the solutions. Brainstorming is an essential tool to generate multiple ideas, to interlace different mindsets and visions, and to consider strengths and weaknesses of the proposed possibilities. Such prototypes are low-fidelity visual models of the proposed solution that serve multiple purposes. Firstly, individuals can embed their early ideas into a tangible representation to clarify, explore and communicate them. Prototypes also allow to bridge among people with different experiences by converging on a common ground [112]. Secondly, since prototypes do not need to be fully functional or polished, their easy implementation requires a scarce investment in terms of time and financial resources, resulting in the generation of more than one single solution. More importantly, each promising solution can receive early feedback from users and other stakeholders that can be used to gradually refine it. If there are changes to be made, or alternative ideas to consider and compare, receiving criticism before the actual development of the artefact is a rather efficient strategy. Building and testing should almost be synchronous: it is good practice to seek constant feedback through continuous testing to refine the proposed idea. Thus, the generation phase must be conceived rather in terms of iterative cycles than in terms of a straight line that brings from a theoretical analysis of the problem to one applied solution.

4.2.6.3 Development and Evaluation

When one solution has been chosen, and consequently implemented, than it must be evaluated to determine whether it is able to accomplish its intended goals and whether the initial hypotheses are confirmed or negated. Design assessment measures concern the people and their interaction with an artefact, what is generally known as user testing: the usability of the artefact and the user experience (UX) of the people. Usability is considered the ability of the user to use an artefact to carry out a task successfully, whereas the user experience is concerned more broadly with a user's interaction with the artefact, comprising her thoughts, feelings, perceptions, and expectations [19]. Metrics measuring usability, e.g. task success or error rate, and user experience, e.g. subjective user's satisfaction, are considered product and technology neutral and can be applied to different contexts. Data can be gathered, for instance, through user interviews, questionnaires, or think aloud protocols, where the user is asked to verbally articulate her thoughts while performing a task [14]. On the bases of the observations about how the users accomplish tasks, the final design of the product is vetted.

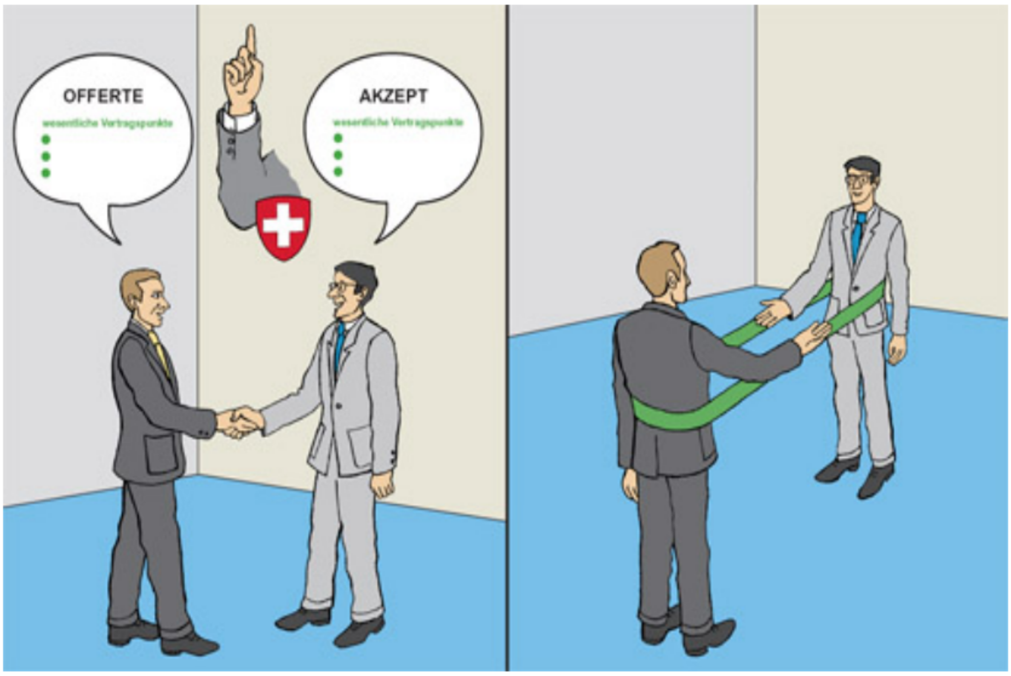
4.3 Legal Visualizations

Legal design can be considered the natural evolution of the incorporation of visualizations in legal texts [42] to make legal communication more understandable and user-friendly. Not only plain language, but also visual elements can greatly contribute to enhance the comprehensibility of the law. Successful communication is not only based on a careful choice of wording, but also on the organization of the information and on the means to transmit it, that must not be exclusively verbal but can also rely on other channels, like the visual one. Visual structure and images are the most commonly used visual tools, but many different graphical cues can be employed, as this Section attempts to demonstrate.

Legal design sets its roots in the *Rechtsvisualisierung*'s (visualization of the law) movement, originated in the early 2000s in German-speaking countries. In her seminal work, Brunschwig [52] codified norms from the Swiss Civil Code as comic-like drawings (*see* Figure 4.1), based on some user-centered guidelines. What is more important, she sets the foundations of a science of interpretation of legal images. She is therefore the first scholar to compare lawyers and designers, suggesting the idea of legal designers[5].

Knowledge visualization (*see* also Section 3.3) designates "the use of visual representations to improve the creation and transfer of knowledge be-

[5] For more on the similarities between lawyers and designers, *see* [135, 131, 136]



OR Article 1 Section 1

For a contract to be concluded, a manifestation of the parties' mutual assent is required.

Figure 4.1: Example of visualization of the conclusion of the conclusion of a contract through mutual assent [52]. Retrieved from http://www.rwi.uzh.ch/en/oe/zrf/abtrv.html

tween at least two people" [90, p.3] to enable individuals to express themselves with richer means and to re-construct, remember and apply such knowledge correctly [90]. It is not only a matter of transfer of knowledge: it concerns how recipients acquire that knowledge, but also how they use it (*see* Section 4.1). For such reason, the party responsible to initiate the transfer of knowledge must attentively find a way that will enable the other party to use it. Adequate knowledge visualization can support learning, reasoning and memorization [90] and fight information overload. Later in this section, the attested benefits of using visual cues in combination with texts to enhance the comprehension of legal matters will be presented.

4.3.1 Reasons

The studies presented in the next paragraphs consider legal documents as artefacts, namely objects built or shaped by the human being to achieve a specific purpose. Artefacts can facilitate or, on the contrary, hinder human learning[6] and action [231] (*see* also Section 4.1). Optimal learning happens when information is structured and presented in a way that suits best human cognition. Cognitive overload (*see* [273, 274]) happens when a significant amount of the working memory's resources are occupied, thus impeding higher-order process, such as schema building and inference making, that are necessary activity for learning and understanding. Indeed, comprehension errors, slower reading, and slow task completion are classic manifestations of cognitive overload. The visual support to text is also motivated by theories about multiple channel processing [218]: the cognitive processing of visual and verbal information is distinct and one modality can integrate and reinforce the other, especially in terms of memorability [124]. Finally, in order to be learned, information must be presented in a way that facilitates search and integration. The role of an artefact's is not determined by what it is, but rather by what it does [231]: taking the example of a privacy policy, both an impenetrable lengthy document and a short visually structured document in

[6] The term learning here is used to indicate any acquisition of new knowledge or skill

principle attempt to reach the same goal, but they are differently designed and facilitate the human being's learning and action differently.

4.3.2 Attested Benefits

A few researchers have experimented with a variety of graphical devices that can make legal information more meaningful and more easily processable by legal experts and non-experts alike. Much research that will be introduced in the following clearly suggests that the visualization of information can unburden the cognitive load derived from reading and understanding complex documents, such as legal texts. Text and images can be treated in a complementary manner, where each represents what the other can not convey. Visualization can even lower the chances of misunderstandings, that could give rise to litigations, because it elicits information by making abstract concepts easier to grasp. For instance, [132, 232] argue that a timeline displaying a contract termination clause could have prevented a considerable lawsuit over the meaning of the clause.

Thus, visual elements of various kinds are entering the text-oriented realm of the law, which in the modern age is characterized by the total absence of graphics, with few exceptions, such as the highway code (*see* also [289]) and patents [46]. Similarly to the Directive on Consumer Rights [227], even the GDPR recommends icons to support the cognitive effort derived from the navigation of lengthy and cumbersome privacy statements. As already highlighted in [254], this is arguably the first explicit acknowledgement of the potential benefits of visualizations in the history of data protection law.

There are several different visual representation techniques, depending on the type of information that is represented, on its goal, on its intended audience, on the context where the visualization will be placed, etc.. In the following, a variety of examples of visualizations of legal content with different functions will be provided.

4.3.3 Educational and Explanatory Purposes

Visualization for legal matters has been extensively used for educational and illustrative purposes. For instance Law for Me[7] and LawToons[8] aim to enhance the understanding of rights and legal processes through cartoons for people without legal background. Initiatives in this sense are countless, and even includes the common visualizations of rules aboard of public transport.

Other initiatives take the shape of guides that do not replace the law, but translate into a more easily graspable manner the content of the law for specific audiences. The Center for Urban pedagogy in NYC, for example, realized a guide to the juvenile justice. System to explain to under-aged individuals through comics, timelines and flowcharts their rights when they get arrested [110]. One illustrative example of the benefits offered by the visualization of rules arising by demonstrated people's needs and that enforce users' rights through clearer (visual) communication is "Vendor Power" [111]. It is a guide that visualizes the complex New York's code that governs rules and rights of the city's street vendors. Although it is especially conceived for the non-native English speakers, but it has also proven useful for those police officers that made inconstitent fines according to a wrong understanding of the laws.

However, the visualization of legal information for educational purposes has different scope, goal and audience with respect to visual elements used in legal documents that contain enforceable rules.

4.3.4 Visualizations in Contracts

The private legal practice has seen the most intensive experimentation about visualizations, especially for what concerns contract visualization (*see* the seminal work of [231] for a complete bibliography). Contract visualization is not only enjoying success as managerial practice[9]), but also as a scholarly

[7] http://lawforme.in/

[8] http://www.lawtoons.in/

[9] *see* for instance https://www.visualcontracts.eu/

exploration.

With contracts growing in size and complexity, a "paradigm shift" [130]that builds on the Proactive Law's principles (*see* Section 4.1.1) and introduces innovation in the contracts' formats has also slowly started to happen. "Contracts contain vital business and relationship information, not just legal provisions, [...] that need to be translated into action" and this is why "contracts need to be *designed*, not just drafted" [137, p.375]. Hence, visual elements have started to appear in contracts with the goal of supplementing the text and enhance contract searchability, readability and usability [138, 231] (*see* e.g. Figures 4.7 and 4.8). When visual communication complements textual communication, the strengths of both can be leveraged to create better legal communication [231].

The visualisation of contractual clauses clarifies the meaning and thereby supports comprehension between parties, e..g about their rights and responsibilities. Empirical evidence suggests that parties understand the contract's content faster and more accurately when it includes visualizations [229, 233, 236], compared to a text-only contract. In the online context (e.g. end user license agreements), the inclusion of visualisation increases reader's attention and time that people devote to reading online contracts [173, 174] and improve comprehension accuracy [49]. Such interventions are accomplished at the level of information architecture, where a clear information hierarchy is provided, with the help of elements such as icons and comic-like vignettes. These elements can reduce cognitive load derived by lengthy agreements and assist the activity of skimming through the text to find relevant information [173, 234, 231]. They can also support memorization and recall of legal concepts [233]. Spatial and temporal contiguity of text and related images support learning and retention of information more than if the two elements are presented separately [231].

Not only the usability, but also the user experience is positively impacted by the inclusion of graphical elements in legal documents, that are thereby perceived as more pleasant to use [236], while the party that drafted it is perceived as more trustworthy, respectful and collaborative [229], due to her effort to communicate more clearly [173, 174]. Moreover, visualizations can even support the activity of those drafting the documents and bring clarity and certainty to the drafting [42]. Visualization assumes particular relevance for contract parties that have low literacy levels or low language proficiency [137, 49], as the next examples will show.

Not only can visualizations be inserted into contracts, but a visual narrative can even completely replace the text of the contract instead of merely supporting it, with no other text overriding the visual representation [137]. In 2016, the world's first legally-binding[10] comic contract[11] was adopted by a fruit-picking South-African company. Their seasonal workers have low literacy levels, frequently speak other first languages than English, and are not well informed of their rights and duties with traditional written contracts. For instance, workers have the right to sick leave, but if they do not know that they have to communicate such event and fail to arrive at work, they are disciplined. However, if such a situation is visually explained (*see* Fig.4.2), both parties understand and are satisfied. In addition, in some cases visualisations can even be more expressive than words because they allow to present the contractual terms contextually, i.e. in a specific situational and temporal setting, and suggest the tone of the relationship (friendly, courteous, formal) [137].

Comic-form contracts have also been used for Non-Disclosure Agreements and Intellectual Property agreements [175], whereas the Australian firm Aurecon has designed and adopted legally-binding visual employment contract, by eliminating more than 4000 words [47]. The company and its employees are represented as comic-based characters, to convey the culture of innovation at the basis of the company's culture and to make contracts accessible

[10] A court decision has yet to come. But the attorney author of the comic contract, on the basis of some experienced lawyers' comments, is confident that it will be legally binding at least in all the common law countries. The creators of Aurecon, together with a former Chief Justice of Australia, are positive about the enforceability of comic contracts [47].

[11] *see* [79] for a preview of the comic contract.

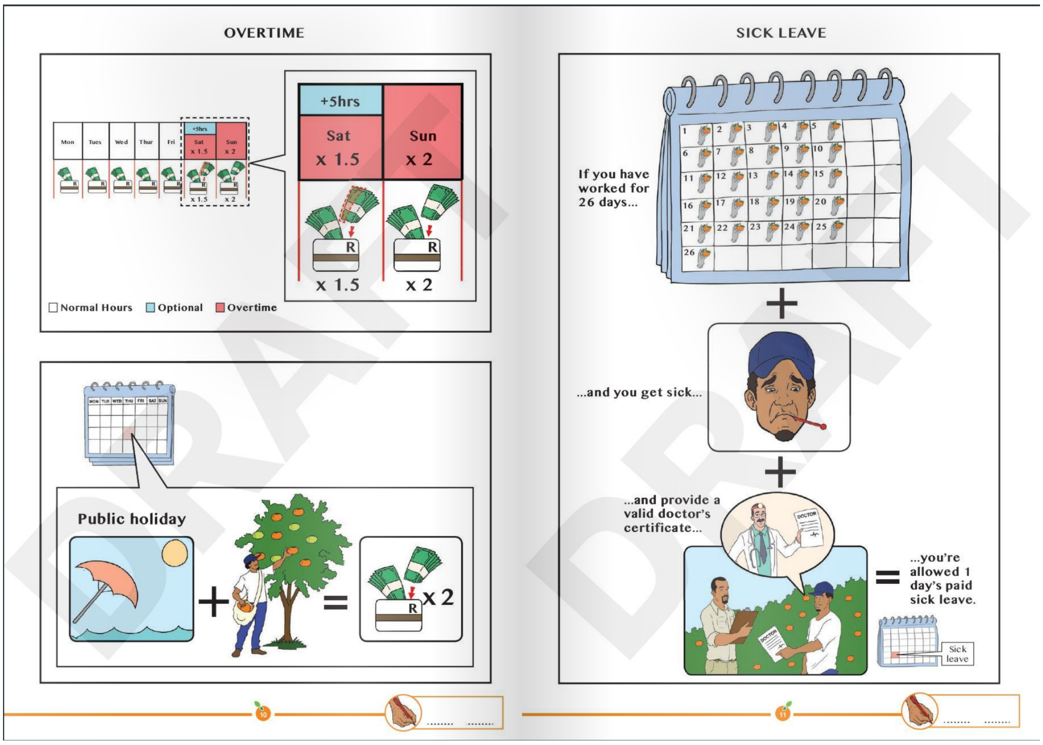


Figure 4.2: A comic contract's page displaying the right to sick leave and its conditions. Reproduced with permissions [79]

by its multicultural employees, by also conveying open and transparent relationships.

4.4 Law, Communication, Semiotics, and Design

4.4.1 A Communicative Perspective on the Law

In [255], we have introduced an assumption that constitutes the foundation of the present project: law can be analysed from a communicative perspective [285], thus parallelisms between communication theories, visual legal communication, and design can be drawn. "Human action implies interpersonal relations and, thus, communication. As a consequence, if law offers a framework for human action, it also offers a framework for human communication" [285, p. 7]. Although an analysis of law under the communicative perspective is outside of the scope of this work, it is fundamental to acknowledge the bidirectional dimension of legal communication. Such bi-directionality is shared with any kind of human communication. Moreover, "law itself is also essentially based on communication: communication between legislators and citizens, between courts and litigants, between the legislator and the judiciary, communication between contracting parties, communication within a trial" [285, p. 7]. It is this assumption that provides a framework of analysis for the present project.

4.4.2 Models of Communication

There exists a number of definitions of communication. Two definitions seem relevant for the present research: communication as "intentional transmission of information by means of some established signalling system" [191, p. 32] and communication as "the practice of producing and negotiating meanings under specific social, cultural and political conditions" [262, p. 8]. The nature of the communicative process (represented in Fig 4.3) can be

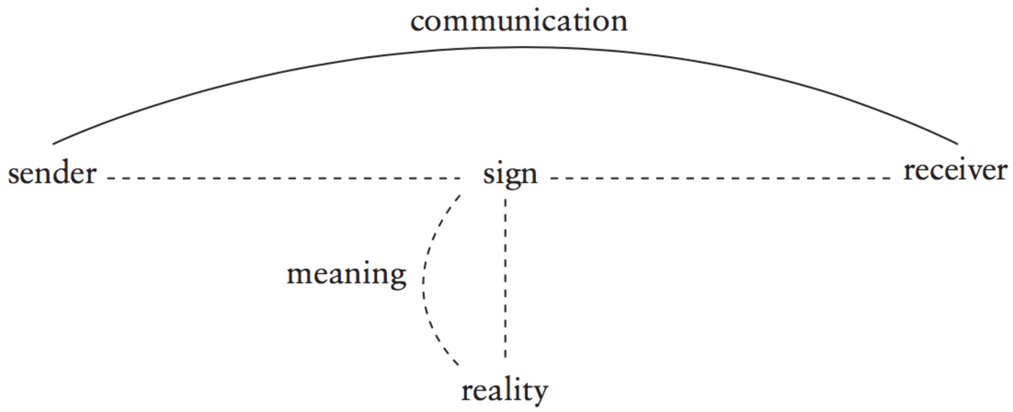


Figure 4.3: A simplified model of communication. Adapted from [285]

applied to any type of communication. Firstly, for its interactional nature, communication is a process that takes place at least between two parties: on the one end of the model stands the sender (the addresser), who encodes a certain meaning in a message and addresses it to a receiver (the addressee)[12].

The model of communication shown in Fig. 4.3 must be understood in combination with what is known as semiotic triangle, i.e. the triangle of meaning. According to Peirce [237], the father of semiotics, a sign comes from a triadic relation among three components: the representamen or symbol (i.e. that which represents, e.g. a word, an interface symbol or an icon), the object or referent (i.e. that which is represented, e.g. a concept), and its mental interpretant (i.e. the process of interpretation) [124] (*see* Fig. 4.4).

The meaning is “a content which the sender has given to the [symbol] in order to give some message about reality” (i.e. the process of representation), but it can also be “the result of the interpretation of the [symbol] by its receiver” [285, p. 128] (i.e. the process of interpretation). In other words, the two possible descriptions of meaning do not necessarily coincide because what is meant by the sender (i.e. the sender-meaning [285]) does not always correspond to the addressee’s interpretation (i.e. the receiver-meaning [285]). The effectiveness of representation depends on what is represented,

[12] For a fundamental theory of communication, *see* [170]

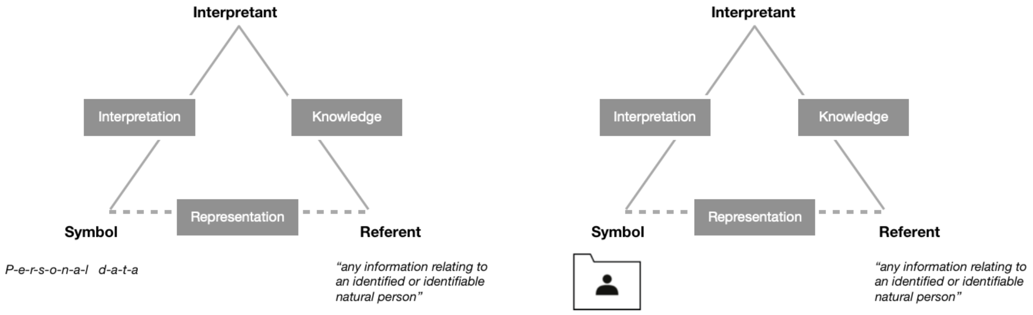


Figure 4.4: Two semiotic triangles where the same referent, i.e. the concept of personal data, is represented by two different symbols. On the left, the symbol is a written, linguistic utterance: note that is the English spelling in Latin characters, but it could well be in a different language with different alphabet (e.g. Greek, Cyrillic, Arabic, etc.). On the right, the same referent is represented by a graphical symbol: an icon.

on how it is represented and on the encoder. Interpretation is the process of understanding the meaning of a sign [124]. The activity of sense-making does not happen objectively, but it is rather dependent on the interpretant, her intrinsic characteristics, and, thus, her mental models, which is the way in which the sign object is recalled (*see* Fig. 4.5). In other words, different interpretants could understand the same sign differently, because their specific background and past experiences influence the process of interpretation. This is why in any kind of communication, included in legal and visual communication, it is crucial to consider all the parties involved in the communicative exchange.

There is not only a subjective dimension to meaning, but also a socially- and culturally-determined dimension. The interpretation of meaning (i.e. the relationship between symbol and interpretant) also depends on the knowledge of a social and cultural system of signification. For instance, an English speaker would be able to interpret the linguistic utterance in the Fig. 4.4, but she would not be able to interpret the corresponding Greek translation without having knowledge of this language and its alphabet. Moreover, the meaning of "personal data" also depends from the culture, and hereby the legal framework, where it is set: in the European Union, it might have a

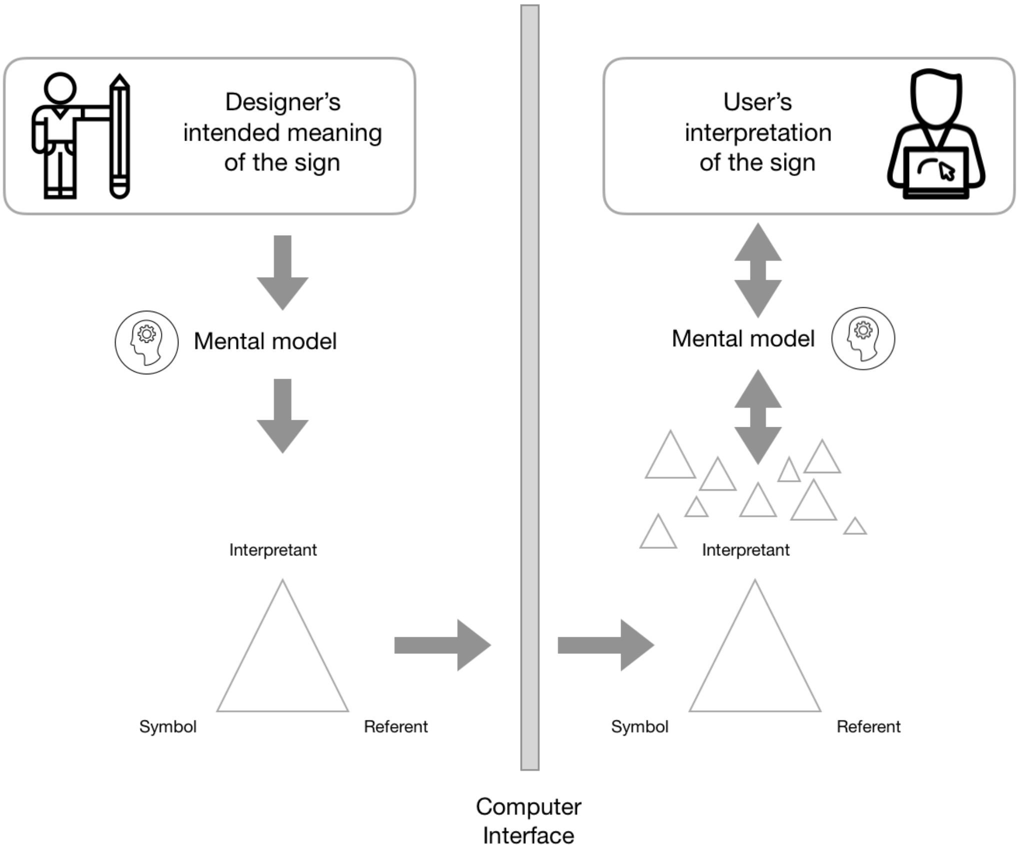


Figure 4.5: This image represents the possible discrepancies between the meaning intended by the designer and the meaning interpreted by the end-users. For the communication to work smoothly, the two meanings should coincide. Adapted from [164]

meaning that is slightly different from another legal culture. Also the interpretation of graphical symbols is entrenched in a cultural system of signification: for instance colors may assume different nuances of meanings according to the country, but also icons give rise to different interpretations if they are not indexical and not standardized (this point will be explored in Sect. 6.1). This is why, some scholars [156] have added a supplementary dimension that gives meaning to the semiotic triangle for an icon: that of context. Chapter 6 will discuss the importance of context for the correct interpretation of icons.

4.4.3 Design as Communication

"[O]ne of the principal functions of design is to communicate" [148, p. 27], i.e. to provide signals to people, for instance about how a technology functions or about what actions are permissible or recommended in a certain interface: "[t]he way something is built communicates information to people about how it works" [148, p. 140] . In general, "well-designed objects are easy to interpret and understand. They contain visible clues to their operation. Poorly designed objects can be difficult and frustrating to use. They provide no clues - or sometimes false cues. They trap the user and thwart the normal process of interpretation and understanding" [212, p. 2]. It becomes evident how these observations can be applied to information design and interface design.

Within this view, design can be understood as a conversation between designer and user: as such it is not mono-, but rather bidirectional and usually takes place when the designer exits the scene [80]. Hence, the model of communication is not only important for interpretative theories of law,[13], but also for the present investigation about and around the role of design. Design can be considered a communicative process [80, 255, 165, 74]: the designer attempts to encode a certain meaning in an artefact (such as an icon or a graphical user interface) so that the end-user will understand that

[13] Such a vision is outside of the scope of the present project, thus it will not be examined. But *see* e.g. [285].

intended meaning (such as an icon's function) and act consequently without frustration nor errors.

For instance, the iconic representation of a printer on a button in a editing software suggests to the user the idea that it is possible to click on the icon to achieve the goal of printing out the document. Of course, as it will be highlighted multiple times in Chapter 6, there are a number of factors at stake so that this transaction works smoothly: for example, the printer on the button must be recognizable and be based on a shared visual convention, whereas user's previous experience with the editor or with a similar software creates a framework of action on which the user can rely to infer the function of the button.

Despite the correspondence, it is "the existence of expressive intent and interpretative response" that delineates the framework to treat design within a communicative perspective and the design products as the message or medium of a sender-receiver process [74, p. 425-427]. Thus, design is a form of mediated communication, like written communication: the interpretation of the message embedded in the artefact (e.g. icon, button, visualization) is carried out in a different time and place than its production. The end-user must interpret the artefact without direct access to the designer's intention. This means that there cannot be the meaning negotiation activity that takes place in instantaneous communication, if the intended message is not grasped by the addressee. It is therefore crucial to design a symbolic system that users will easily interpret: the sender must anticipate possible misunderstandings and craft a message that will need to be decoded by different audiences without receiving any feedback by its encoder.

Thus, "a good correspondence between intention and interpretation might be considered a requirement for design success" [74, p. 429]. This is why usability is a key dimension: as a discipline, it studies how to design artefacts that can be used by individuals to achieve their goals with the least possible effort and provides measures to determine the effectiveness of certain design choices. Similarly, legal design tries to make legal documents and legal sys-

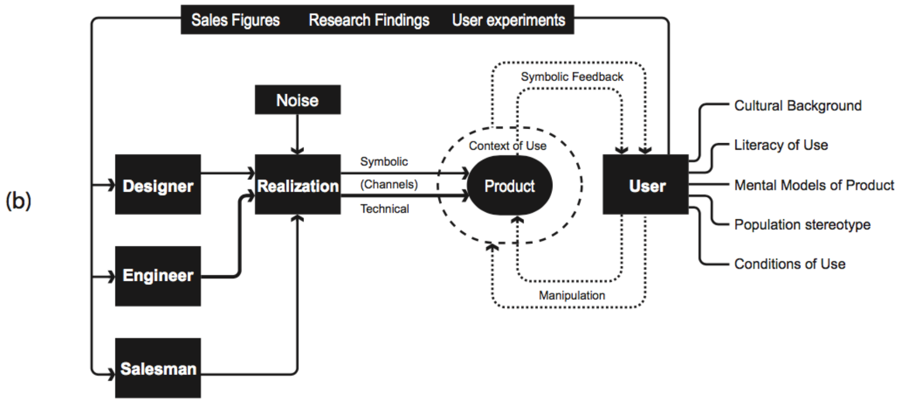


Figure 4.6: Semantics in the design and use of artefacts [180]

tems more usable. Anticipating the factors that will lead to problematic interpretations is therefore fundamental: it is the interpretation rather than the intention that determines success of use [74]. There exists many models for design as communication, for instance the one in Fig. 4.6. This draws a generally applicable model that includes the context of use of a certain artefact, but also users' characteristics that influence the sense making activity (such as the cultural background, the literacy of use, mental models of the product)[14].

Visualizations are forms of communication similar to design choices and words, thus it is fundamental to make their meaning the more accessible as possible, for example by relying on existing shared visual vocabulary. It would not be reasonable, for instance, to use a known symbol (e.g. a scale) but assign to it a different meaning from its conventional use (e.g. "peace" instead of "justice"). If wrongly interpreted, the visual representation of a legal concept would create obscurity in lieu of transparency [255]. This can happen when there is a mismatch between designers' intentions and users' interpretations. Some concrete examples will be provided in Chapter 6. Design can also communicate certain permissible or recommended actions

[14] For more on the relationship between semantics and design, *see* [179]

to the users, but with deceptive intentions (*see* Section 5.3). “In digital services, design of user interfaces is in many ways even more important than the words used” [9, p. 7] to direct individuals towards intended behaviors. In conclusion, design can be considered as a form of communication and can even be more effective than verbal utterances.

4.5 (Mis)interpretation of Legal Visualizations

The law is verbocentric [120] and has been traditionally based on words and texts: “[l]aw, like most other disciplines or practices that aspire to rationality, has tended to identify that rationality (and hence its virtue) with texts rather than pictures, with reading words rather than ‘reading’ pictures, to the point that it is often thought that thinking in words is the only kind of thinking there is” [105, p. 4]. Indeed, “[l]aw is language” (for instance [120]): legal literacy is based on reading and writing legal texts, and assigns a fundamental role to verbal rhetoric [120], whereas legal norms are extensively treated as linguistic utterances (on the incompleteness of such approach, *see* e.g. [85]).

Nevertheless, a visual turn in the legal world [263, 105] has been auspicated also as a result of a general change happening in the (digital) society at large: with the widespread use of GUIs, individuals are becoming more and more familiar with the use of images and icons to search, navigate and make sense of any type of information. People have learned to simultaneously visit multiple sources to extract the more meaningful and relevant information for their inquiries in a limited amount of time [39]. Besides, human being do not make use of one channel at a time to interact with the outside world: they are rather “multisensory beings” that live “in a multisensory world” [120, p. 231] and as such generate and decode messages by using multiple sensory channels that simultaneously involve vision, hearing, and movement at least (*see* also the dual code theory in Section 4.3).

Images are becoming the main cultural medium in the modern society

[46] and, even in the legal world, evident signs of this legal-visual revolution [120] are emerging: legal norm images (*see* [52]), visual jurisprudence (*see* [263, 105]), legal visualizations for educational purposes (e.g. *see* Sect. 4.3.3) and legal visualisations in private legal practice (*see* Sect. 4.3.4).

With the entrance of graphical elements in the law, some scholars have started to speculate about the process of their interpretation and to worry about the possibilities of their misinterpretation. Apart from the visual elements used before the modern age to convey legal messages to illiterate people, the law in the contemporary age is verbal, hence it has developed instruments and methods to encode legal concepts in words and to decode them accordingly. However, a debate around visual jurisprudence is emerging (*see* for instance [201, 209, 245, 268]), whereas iconographical and iconological methods have been compared with legal hermeneutics [46].

Some considerations should be made to counterweight the criticism moved against visual legal communication. Firstly, contrary to what is commonly believed, the majority of visual elements that are inserted in legal documents accompany the text and do not aim to replace it [91]. Thus, the priority of the written language is safeguarded, whereas graphics mainly have the aim of illustrate and clarify legal terms and actions, focus attention on important clauses, support readers to easily find information, among the possible functions [39]. Even in the case of comic contracts, where graphics dominate the written word, a combination of verbal and visual elements is necessary to ensure the conveyance of the right message. As recalled earlier (*see* Section 4.3), words can convey what visuals cannot convey and vice versa.

Moreover, for instance in contracts, visualizations have been used as tool to create value in the contractual relationship to "transform contracts from traditional legal instruments for rent-seeking of risk-shifting toward [...] devices to facilitate better collaboration, relation-building, innovation, strategic planning, and social value" to ultimately achieve "better understanding, communication, and trust" [39, p. 49], in a classical proactive law perspective (*see* Sect. 4.1.1). If legal documents should be used as artefacts to achieve specific goals, not only by judges and lawyers, but also and foremost by people without legal expertise, a major transformation in the conception of legal documents should occur. Instead of thinking of such documents from a purely legal perspective, they could be reconsidered in a functional perspective, as tools that enable the communication between parties and that ensure comprehension between them, instead of focusing on the prevalence of one party over the other in case of a dispute. Instead of considering privacy policies as a liability coverage, such documents could deliver value, increase trust, and seek win-win opportunities for both sides. Of course, visualizations cannot solve unfair terms or unbalanced relationships, but can otherwise promote a culture of transparency that seems much needed, but also legally mandated, in the privacy modern world.

In this light, then, visual aids have the aim to support comprehension of the respective roles and responsibilities, ultimately seeking reciprocal understanding, with the very aim of eliminating or at least lowering the possibility of dispute arising. Visualizations can be even more helpful when the parties that must understand a binding document are not legal professionals: in such case, visualizations that improve mutual understanding could influence the perception of the law "as less of a power game, manipulated by forces beyond ordinary people's control, and more as a framework for individual and collective flourishing" [39, p.50].

Linguistic and non-linguistic communication have different features and therefore can also achieve different goals. Images, as a different communication channel, cannot completely express what the language of the law can convey, for instance in terms of legal or technical nuances. On the other hand, visual elements can be more expressive than words: since the visual dimension pre-dates language, it is also associated to an emotional and not only intellectual dimension - this is why, for instance, comic contracts can also convey the tone and feeling of a relation (*see* Section 4.3.4), while colors can be shown but not described through words [85]. Indeed, visual meaning making is different than verbal meaning making [263]. In certain cases, graphic signs express legal concepts even better than words, which means that it would be possible, but non-functional [85], to use linguistic utterances at their place: for instance, in the case of traffic signs that do not complement, but rather constitute an integral part of legal provisions.

All in all, although in [255] parallelisms between communication, design and legal hermeneutics have been explored, methods of interpretation of the visualization of the law fall outside of the scope of the present dissertation. This work is rather conceived in a functional perspective and from a layperson-centered point of view: as non-experts are not expected to have the necessary knowledge and skills to use hermeneutical tools to make sense of the legal information they encounter, the same expectations hold for the interpretation of legal visualizations.

Finally, legal communication is not generally tailored to respond to the needs of its end-users, mostly non lawyers. Notwithstanding the research, movements and regulatory actions that have proposed ameliorations, this reality has not undergone much change. Legal communication is mainly not human-readable: it is utmost machine-readable. However, there is no outcry from legal professionals on the risks of misunderstanding the traditional legal communication. Hence, the criticism on the use of images, if based on the argument that legal visualizations risk misinterpretation, appears weak. On the contrary, much research shows that visualizations facilitate comprehension of complex legal terms, whereas traditional legal texts make this task excessively difficult.

In the next few years, while legal visualizations will spread, there will be the chance to find out whether visual documents will be object of a litigation will. Only in that moment it will be possible to determine how they will be considered and interpreted by the jurisprudence. More importantly, it would prove beneficial to compare the number of cases that end up in court because their textual provisions have been misunderstood with the number of cases whose visual terms have been incorrectly interpreted. Besides , the adoption of a proactive approach attempts to avoid courts altogether because reciprocal understanding is highly valued and actively sought. In this perspective, legal documents are crafted for the parties that will use or implement them, and not for the judges.

Nevertheless, risks of misinterpretation do exist and some examples concerning data protection icons will be provided in Chapter 6. Possible remedies proposed in this work are participatory design methods and, especially, user studies that evaluate the effectiveness of the visualizations, that can point to major flaws and allow for clearer communication. However, different sorts of legal visualizations exist, each entailing a different level of risk of misunderstanding. Indeed, no sense-making activity can be error-free, not even the interpretation of the written word, but it is true that some visual devices might cause more doubts and uncertainty than others. For instance, indexical graphical elements closely resemble the entity they are meant to depict and images' indexical qualities play a major role in smooth and intuitive communication. Other graphics attempt to represent less tangible or intangible ideas through metaphors or learned conventions: in these cases, it is suggestable that the word prevails [39], as it will be argued in Section 6.1.

In conclusion, no certainty can be granted for a correct interpretation of visualizations. But this is not an exclusive problem of the law. One incentive to ease correct sense-making is the use and re-use of design patterns, that are either based on shared mental models or that create new mental models that can be leveraged for future encounters with similar instances, because the human brain has an innate ability to recognize patterns.

4.6 Legal Design Patterns

Patterns are re-usable forms of a solution to a commonly occurring problem [136]. The original idea of patterns stems from [21], who collected re-usable architecture and design solutions for other architects. The idea was later applied to software design and gained widespread acceptance with [117]. From that moment, design patterns have been used in various fields, from computer science, to interface design, to privacy and security (for these, *see* next Chapter). Patterns offer the benefits to extend communication among people working on similar problems, to exchange ideas and knowledge among people working in different domains but on similar challenges, and also to discuss and possibly set standards and best practices on best solutions for a given problem [134, 137].

The development and re-use of recognizable patterns is relevant not only for those people that need to find a solution, but also for those people that will need to make sense of it: the interpretants can apply their previously acquired knowledge to a novel artefact and understand how to interact with it with lower transaction costs. For instance, a number of patterns are used consistently in interface design: e.g. a button on the right of a widget's window, when clicked, lets the user navigate to the next page. Colored buttons are active buttons and thus offer the affordance to be clicked, whereas greyed-out buttons are inactive. Users do not need to explore, learn and remember a new language every time that they encounter a new element, but they can rather rely on a learned pattern language and apply the previously acquired knowledge to new instances of that pattern: for example, once that an individual has learned how to make sense of the different elements of a diagram, she can resort to this acquired mental model to interact with all the future diagrams she will encounter.

Indeed, although patterns are general containers and remain on an abstract level, allowing for multiple practical implementations, they also "retain distinguishing features that allow us to recognize and re-create [them]" [136, p. 3]. For their generic nature, they are perfect candidates to frame visualizations in a non-prescriptive manner [136], as legal design as a discipline tries to attempt more broadly, and to provide solutions that must be considered in context (e.g. according to users' characteristics) (*see earlier*). For their very nature, patterns are, hence, strictly correlated to their constant and widespread use and reuse: their application in practice determines their function and their replicability.

4.6.1 Contract Design Patterns

As introduced before, contracts can be considered as information artefacts that communicate rights and responsibilities to the two parties, so they can be conceived as a result of information design. As such, they must not only be conceived merely as textual content, but as artefacts that can be used with success by the two parties.

Patterns have started to emerge conspicuously in the field of contract design and contract visualization, although they appear only in a small percentage of contracts worldwide. Nevertheless, it can be safely assumed that contract patterns exist and are gaining acceptance - indeed, they have been proposed as solutions to properly present, communicate and apply contracts [133]. A pattern design library [143] that organizes the main contract design patterns according to four categories has been published. Process patterns concern the act of crafting the agreement, whereas layout composition, visualization, and clause text are about the crafting of the document itself.

Other contract patterns have been developed in terms of function they serve [291]: patterns that support strategic reading (e.g. skimmable headings, alert icons), patterns that support explanation (e.g. timelines, exemplars, layered explanations), patterns that support an effective user response (e.g. checklist), and patterns that support reader engagement (e.g. topic icons that break up the long text and highlight particular information).

Further research [136] focuses on contract visualization patterns and proposes a more fine-grained classification for the patterns inside the categories:

1. Visual organization and structuring patterns “organize and structure texts visually by means of layout, page design, and typography in order to increase readability and legibility and support strategic reading activities such as searching, skimming, and selecting relevant content” [136, p.12]

2. Multimodal document patterns transform the document into a visual format where text and images are fully integrated, e.g. comic-based

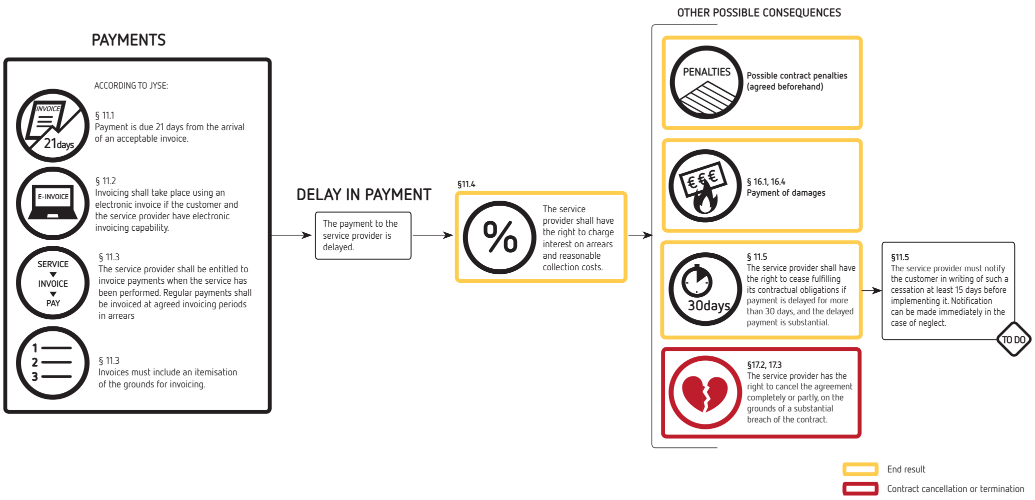
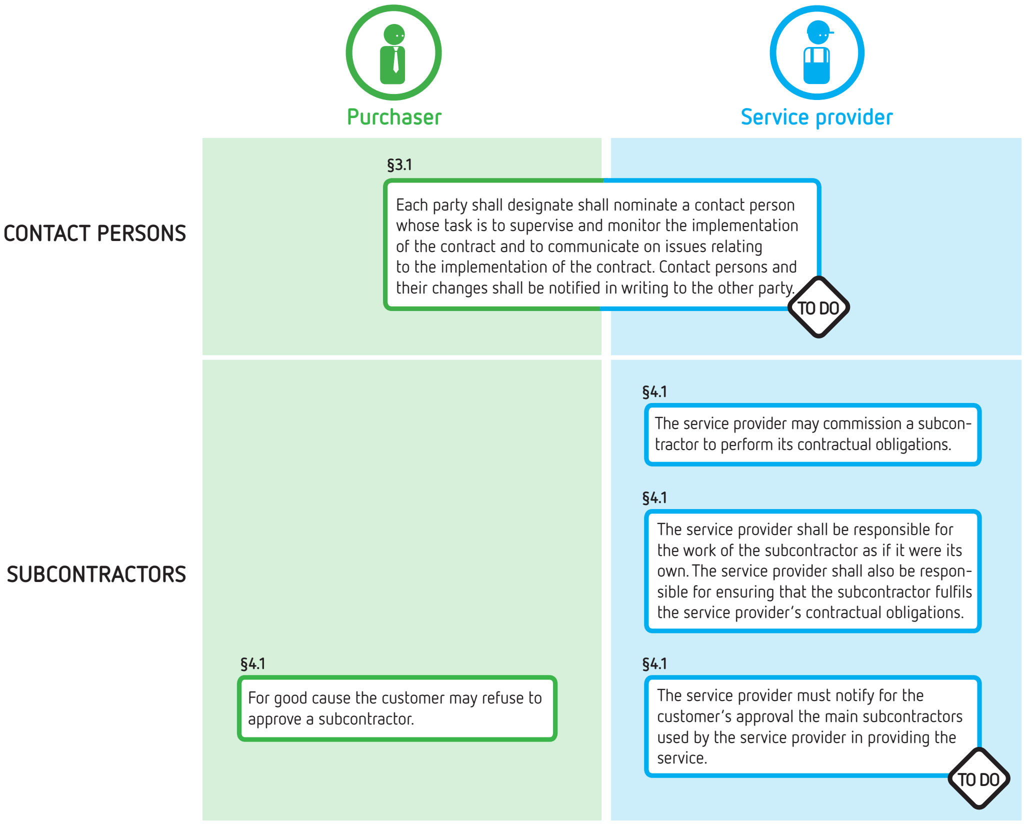

PAYMENTS
ACCORDING TO JYSE:
INVOICE
21days
§ 11.1
Payment is due 21 days from the arrival of an acceptable invoice.
E-INVOICE
§ 11.2
Invoicing shall take place using an electronic invoice if the customer and the service provider have electronic invoicing capability.
SERVICE
INVOICE
PAY
§ 11.3
The service provider shall be entitled to invoice payments when the service has been performed. Regular payments shall be invoiced at agreed invoicing periods in arrears
1
2
3
§ 11.3
Invoices must include an itemisation of the grounds for invoicing.
DELAY IN PAYMENT
The payment to the service provider is delayed.
§11.4
%
The service provider shall have the right to charge interest on arrears and reasonable collection costs.
OTHER POSSIBLE CONSEQUENCES
PENALTIES
Possible contract penalties (agreed beforehand)
€€€
§ 16.1, 16.4
Payment of damages
30days
§ 11.5
The service provider shall have the right to cease fulfilling its contractual obligations if payment is delayed for more than 30 days, and the delayed payment is substantial.
§11.5
The service provider must notify the customer in writing of such a cessation at least 15 days before implementing it. Notification can be made immediately in the case of neglect.
TO DO
§17.2, 17.3
The service provider has the right to cancel the agreement completely or partly, on the grounds of a substantial breach of the contract.
End result
Contract cancellation or termination

Figure 4.7: Example of flowchart used to elicit payment procedures and consequences of delayed payments in the Visual Guide for the Finnish terms of public procurement [232]. ©2013 Aalto University & Kuntaliitto ry. Licensed under CC-BY-ND 3.0.

contracts (*see* Sec. 4.3.4)

3. Visual representation patterns that represent logic, content or prerequisites of the contracts through diagrammatic or pictorial representations; unlike the precedent patterns, the visualizations integrate and disambiguate the text.

This last category includes flowcharts (Fig. 4.7) that express complex conditional structures typical of legal texts and swimlane tables (Fig. 4.8) that highlight vis-a-vis the roles, rights, and responsibilities of different stakeholders [232]. These representations can be generalized to other kinds of legal documents, such as privacy policies, because they visually reproduce logical structures that are typical of these legal texts. As introduced in Section 3.3.4, rights and obligations, as well as conditionals, and consequences of specific actions (e.g. give or withhold consent) can be expressed in a machine-readable language and consequently translated into visual elements. In this light, patterns can be coupled with legal informatics' patterns and made automatically accessible and replicable - as proposed in the present research.



Each party shall designate shall nominate a contact person whose task is to supervise and monitor the implementation of the contract and to communicate on issues relating to the implementation of the contract. Contact persons and their changes shall be notified in writing to the other party.

The service provider may commission a subcontractor to perform its contractual obligations.

The service provider shall be responsible for the work of the subcontractor as if it were its own. The service provider shall also be responsible for ensuring that the subcontractor fulfils the service provider's contractual obligations.

For good cause the customer may refuse to approve a subcontractor.

The service provider must notify for the customer's approval the main subcontractors used by the service provider in providing the service.

Figure 4.8: Example of swimlane table used to illustrate the parties' rights and responsibilities in the Visual Guide for the Finnish terms of public procurement [232]. ©2013 Aalto University & Kuntaliitto ry. Licensed under CC-BY-ND 3.0.

The visual representation pattern that is more relevant for the present research is the companion icon pattern [136], which will be further elaborated in the next chapter to decline it according to the specific research goals of this dissertation. Icons can be used to represent the meaning or the function of the text span they accompany. Such graphical symbols are useful visual devices when the document is long to give salience to certain elements that would be otherwise lost in an undifferentiated text. Thus, their function is to help readers to skim, search the document and identify information quickly and efficiently. Icons are not self-explanatory, since they can not convey subtle and nuanced meanings, but they rather can quickly suggest where a certain information item is to be found in lengthy texts. As the icons' use spread in a coherent way across more documents, they become more and more easily recognizable by individuals. This aspect assumes relevance in the context of patterns: if such visual solutions are to be reused, then it is essential to establish a common and shared language to reduce the initial costs involved in coding and decoding activities of such visual elements. Companion icons is the framework under which the data protection icons that will be introduced in Chapter 6 have to be understood.

4.7 Conclusive remarks

This Chapter has offered an overview over legal design and legal visualization in a communicative perspective. Firstly, in Section 4.1, the notions of legal literacy and document literacy were explored to highlight how comprehension also concerns the ability to act upon certain information. In this design-oriented perspective, legal documents are considered as artefacts that help their users to achieve specific goals with efficiency, effectiveness, and satisfaction (*see* Section 4.2). In order to achieve this goal, legal communication and legal documents should be human-centered, i.e. should be designed with their final user in mind. This is attained with the inclusion of users in the design process, from the analysis of needs and opportunities, to the generation of ideas and prototypes, to their evaluation.

Visual means of communication (Section 4.3), e.g. layout and images, are strategic because empirical evidence demonstrates, that they ease comprehension of complex topics and navigation of legal documents, among the other benefits. A communicative perspective that unites law, semiotics, and design was presented in Section 4.4 and introduced a discussion in Section 4.5 on the risks of misinterpretation of legal visualizations. The empirically-informed creation and evaluation of visual elements for legal matters can reduce the chances of misinterpretation, while replicable and standardized solutions to common problems (i.e. design patterns) leverage on acquired mental models, and not only facilitate the correct interpretation of the visual patterns, but also simplify their generation.

The next Chapter will discuss design patterns for the data protection domain, specifically for what concerns design patterns for transparent communication. Such exploration will be introduced as linchpin of a wider analysis about the role of design for data protection and data protection by design.

Chapter 5

Design in and for Data Protection

In recent years, a growing body of research from various disciplines has focused on the analysis of what determine users' behavior and decisions relating to their privacy, as it was illustrated in Chapter 2. Namely, studies of usability and human-computer interaction have investigated how to make interfaces more usable to enhance people's privacy. Behavioral economics has provided evidence on individuals' actual decision-making process about their privacy online, which is distant from that of rational decision-makers presumed by the law. Other research has examined the design of interface and services to respond to those cognitive biases that are behind disadvantageous or risky privacy decisions.

In the last Chapter, design was considered as a generative, open-ended process that identifies and analyzes existing problems as target areas for the development of new, useful artefacts. This view reflects the work of designers that define user experiences, thus incorporate a contextual understanding of the end-users in the artefact they develop. For engineers, the design of systems assumes a more formal, objective, requirements-oriented activity [148]. Thus, it is clear that the term design has many possible definitions according to the domain of application. In the domain of data protection, it can assume even more nuances: this chapter will examine the meaning of data protection by design and it will describe how the design of communication, interfaces and user experiences can influence individuals' privacy-related behavior. Indeed, a design-oriented perspective is necessary to develop privacy-preserving technologies that intend to achieve privacy-friendly outcomes. Unfortunately, design can also be used for opposite aims, such as creating privacy-corrosive technologies. From a more closely engineering point of view, the (legal) design patterns presented at the end of the chapter will provide a manner to traduce the GDPR's legal requirements related to transparency and consent into practical and re-usable solutions.

The area of privacy and data protection constitute a privileged environment to observe and research the importance that design assumes in people's lives. Data-gathering technology is pervasive in the modern society: its underlying architecture, its function, its interface and its communication, all of these elements directly or indirectly affect data subjects. "Design decisions establish power and authority in a given setting. They influence societal norms and expectations. When people say they use modern information technologies, what they are really doing is responding to the signals and options that the technology gives them. We can only click on the buttons that we are provided. Each design decision reflects an intent as to how an information technology is to function or be used." [148, p. 8]. As the analysis in the last Chapter pointed out, one of the main goals of design is to communicate with users: design predicts their reaction to specific design choices and exert an influence on their behavior to steer it towards desired outcomes. This is why, "[d]esigners and engineers are choice architects" [148, p. 35]: they organize environments were people's actions are guided. Such reflections echo Lessig's words [184] about the power, and consequent responsibility, that code writers have in the shaping of our digital world (the "cyberspace" [p. 2]), decisions from which derive the presence or, conversely, absence of possibilities and opportunities for the users of that digital space. "As the world is now, code writers are increasingly lawmakers." writes Lessig (p. 79) "They determine what the defaults of the Internet will be; whether privacy will be protected; the degree to which anonymity will be allowed; the extent to which access will be guaranteed. They are the ones who set its nature. Their decisions, now made in the interstices of how the Net is coded, define what the Net is".

It is crucial to recall the fact that choice architecture is inevitable [272], as it was maintained in Section 2.4, especially in an online, device-mediated environment: "there is no such thing as a neutral design in privacy, security, or anywhere else" [15, p. 32-33] warn Acquisti and colleagues. The type and order of display of information items on a menu is a choice. The privacy-preserving (or privacy-invasive) default settings of a device, a browser or an app represent a choice. Even the lack of defaults, which causes mandated decisions for the user at a certain moment in time is a choice. The wording used to frame certain risks is a choice, as well as the prominent display of certain information over other information. The lack of information architecture in long legal documents is a choice. Silence or the very lack of choices is also a choice.

"[T]he entire data economy is founded on design that makes tasks easier" [148, p. 29] - or harder. This discussion can be considered in the wider view of data protection by design and data protection by default, presented in Section 5.1: privacy-friendly values can be embedded in the design of a system by default, so that, even without any activity, data subjects are automatically protected. Technologies with the function of promoting transparency (Transparency Enhancing Technologies) acquire a particular relevance in the hereby presented research and can be coupled with (legal) design patterns that aim to achieve better privacy-related communication. This is why, in Section 5.2, a review of the existing panorama of such design patterns is presented, although only a minority of them is commonly implemented and most of them constitute emerging and even futuristic solutions. Each of these patterns is described in a functional perspective and is proposed as solution to one or more problems identified in the first chapter. As repeated several times throughout this dissertation, however, design can not only be used to steer individuals towards privacy-protecting behaviors, but also to steer them towards privacy-eroding choices. In Section 5.3, thus, dark privacy patterns will be introduced.

5.1 Data Protection by Design and by Default

The inclusions of the principles of data protection by design and by default can be enthusiastically greeted as a step towards a behaviorally informed regulation. In other words, the GDPR in this respect considers the limitations of data subjects in the exercise of their rights [20] introduced in the first Chapter and it provides measures that facilitate the adoption of privacy-preserving behaviors. Essential instruments to achieve such goal are data protection by design and data protection by default.

In the past, privacy and data protection have been conceived often as a hindrance to business and technological development or, at best, as an afterthought [63], that occurred after the last stage of implementation of a technology. Data protection has been traditionally perceived as a mere fact of legal compliance, rather than being regarded as a helpful mind-set and operational value to be integrated into an organization's practices [98]. The distance between legal compliance disciplined by lawyers and technological development driven by engineers and business managers, who often find it difficult to translate abstract principles into technical and managerial implementations, started to be bridged in the 1970s with the emerging of Privacy Enhancing Technologies (PETs), a set of technological solutions aimed to minimize the privacy risks because, instead of focusing on *ex post* remedies, they focus on effective *ex ante* protection [148].

This paradigm shift is, thus, characterized by a pro-active attitude, similarly to the attitude assumed by legal design and contract visualization highlighted in the last Chapter: privacy requirements should be embedded into the design and architecture of the system and the business model [98]. Such attitude is at the center of privacy by design, a concept popularized by Ann Cavoukian who broke it down into the seven, now renowned, principles [62]. Privacy by default assumes a central role, not only because it becomes an effective means to attain other compliance goals such as data minimization, but especially because it does not require any active action from the user to protect her privacy. On the contrary, her privacy is automatically (i.e. by default) granted: the inactivity of users does not preclude them by being protected, an inactivity that has been exploited for long time to gather personal data (*see* also Section 5.3). These words echo the behavioral insights analyzed in Chapter 2 and have been taken into account by the GDPR, the most striking example being the explicit ban of pre-ticked boxes for consent.

Even more importantly, these principles explicitely enter the EU legislative framework with Article 25, titled "Data protection by design and by default[1]", which provides that data controllers develop measures to integrate the data protection safeguards from the design stage throughout the whole processing in compliance with the GDPR and the data subjects' rights. Indeed, rather than an afterthought, data processing and relative protections should be the "outcome of a design project" [98, p.6] oriented to achieve the principles set out in Article 5, among which transparency.

Transparency figures among the foundational principles of privacy by design and presumes that any data processing can be understood, reconstructed and, thus, explained[2]. User-centeredness is also a crucial dimension of privacy by design and has the meaning of conceiving and placing the human being, rather than e.g. purely economic considerations, at the center of technology development, while providing her control over her data: privacy-preserving defaults, appropriate notice and empowering user-friendly options are examples of user-centric measures [62].

[1] The term data protection, rather than privacy, is preferred because it identifies the specific obligations that contribute to achieve the more general goal of privacy by design [98].

[2] A discussion about the feasibility of such an approach in the world of Big Data and algorithmic decision-making is outside the scope of this dissertation, but *see* e.g. [228], [53].

One possible approach to implement privacy by design measures is the development of Privacy Enhancing Technologies and the following sections will focus on a subset of such technologies: Transparency Enhancing Technologies (TETs). As ENISA points out [75], any privacy by design methodology must consider the involvement of the user, while designers must address questions about what information must be communicated to users, in what form, and at what time. Legal design patterns for transparency, introduced in Section 5.2, consider these dimensions. In addition, according to the European Data Protection Supervisor [98], digital ethics (in terms of human values and especially dignity) must complement a regulatory approach and guide technological advancements - value sensitive design introduced in Section 5.1.2 is in this respect a crucial instrument.

5.1.1 Transparency Enhancing Tools

Transparency Enhancing Technologies are tools developed with the explicit purpose of diminishing the asymmetry of information between data controller and data subject [304]. They assume particular importance in the modern world scenario because there is a positive correlation between the transparency about data practices shown by a certain organization and the level of trust users and costumers develop in that organization. Any disclosure of personal data is conditional upon a form of trust between the discloser and the recipient of such data [75]. Trust increases the willingness to share personal information and to engage in online shopping activities[3] [275, 301], which is fundamental for the flourishing of democratic societies, but also for the development of the European Digital Single Market and for consequent economic growth.

Research on tools that enable the understanding of privacy and data protection practices is growing: not only because under the GDPR, as illustrated in Section 2.7, transparency becomes a fundamental dimension of

[3] For a review of studies demonstrating the connection between transparency and trust *see* [171].

fairness and accountability; but also because data breaches (e.g. Yahoo) and data scandals (e.g. the Cambridge Analytica scandal) are eroding data subjects' trust in digital services, while raising their awareness and concerns about data practices. Unlike other TETs that focus on the transparency of data sharing and processing practices [171], the research described in these pages mainly revolves around transparency of privacy communication. Thus, it can be ascribed among the tools that provide insights about intended data practices in an accurate and comprehensible manner.

Within the privacy by design philosophy, a privacy policy should be the outcome of a design project, not only for what concerns its content but also its presentation: as discussed in Chapter 1, transparency is not only a matter of substance but also of form. The common attitude around privacy communication has conventionally been reactive rather than pro-active, with endless privacy policies covering any foreseeable eventuality. However, a privacy notice should not be a document confusedly drafted by copying from other similar services and included with the mere objective of legal compliance. It should rather be the natural result of an internal audit about the processing operations carried out by the organization. Moreover, it should not aim at dismissing liability, but it should rather try to prevent problems and be genuinely communicative by reflecting through the transparency of its language and display the transparency of the processing that it describes. Obscurity and vagueness about data collection, processing and sharing can be negatively perceived and create distrust in the service customers and auditors alike, and as such are explicitly forbidden in Article 12 GDPR.

5.1.2 Value-Sensitive Design

Data Protection by design shares many assumptions with Value-Sensitive Design [114], also known as Design for Values [284], an approach to the design of artefacts "that accounts for human values in a principled and comprehensive manner throughout the design process" [114, p. 1186] and "proactively consider[s] human values throughout the process of technology design" [77,

p. 11]. This approach mantains that, while design artefacts are classically evaluated according to dimensions like usability, reliability and correctness, also human values with ethical import (e.g. privacy, security, trust, accountability, transparency, informed consent, fairness, justice, human dignity, well-being, autonomy, etc.) should be taken into consideration as a central design principle.

Indeed, innovation always involves human values [114] and this position challenges the commonly upheld belief that design is a technical, value-neutral task to develop artefacts according to functional requirements [284]. The embedding of values into an artefact defines the affordances and constraints of its users and, hense, shapes their actions, their experiences, and even societies at large. This is why, values should be articulated early on and throughout the whole design process, when there are still relevant possibilities of intervention in the architecture [284]. (Legal) design should therefore be proactive [253], namely possible ethical concerns should be identified *ex ante*, instead of waiting for problems to arise [77].

For example, interface design choices (e.g. the number of clicks required to attain a certain goal) can decrease the ease with which, for instance, users can share personal data on a social network. Such obstacle realized through design choices will also influence the users' decision, not deterministically but at least predictably: users will still be free to pursue their goal, but it will become harder or more time-consuming to do so. Indeed, "[w]e can build, or architect, or code cyberspace to protect values we believe are fundamental, or we can build, or architect, or code cyberspace to allow those values to disappear" [148, p. 81].

5.2 Legal Design Patterns for Privacy

5.2.1 Privacy Design Patterns

Privacy patterns have emerged in conjunction with the privacy by design approach: they translate into practical and concrete engineering solutions the abstract principles of privacy by design. Design patterns offer a description of the problem they aim to solve, they are internally organized in an easy-to-consult standardized template and collected in privacy design pattern libraries[4]. Hoepman [154] devised eight privacy design strategies (then revised by [69]), that are general architectural building blocks to achieve a certain goal on a higher and more general level than design patterns which solve a specified problem. Hence, they constitute “a potential bridge between the legal and the engineering domain” [69, p. 33].

5.2.2 The Inform and the Control Strategy

For the scope of the present investigation, two of such privacy design strategies assume considerable importance: the inform strategy and the control strategy. The first strategy has the goal of adequately informing the data subject when a processing of her personal data takes place and is translated into e.g. effective privacy communication. The second puts the data subject in control of the processing of her personal data and is realized through informed consent (for which the informed strategy is a precondition), or the right to data portability and right to be forgotten. Indeed, these are the sole strategies that presume an interaction between controller and data subject [69]: the controller informs the data subject about the data processing, who is in control of her personal data through the consent given to the controller.

Researchers and practitioners around the world have started to experiment with innovative ways of communicating privacy-related information (*see* for instance [260]), but these attempts are dispersed and often difficult to find or reproduce, thus they do not stimulate widespread adoption and sustainable innovation. This is why, as part of the present research, some of these experiments have been collected [134] in an online legal design pattern library[5]. These patterns, on the model of privacy patterns that aim to

[4] *see* e.g. https://privacypatterns.org/ and https://privacypatterns.eu/

[5] http://www.legaltechdesign.com/communication-design/legal-design-pattern-libraries/privacy-design-pattern-library/.

translate privacy-by-design requirements into practical advice for software engineering, offer practical solutions to the common problems of traditional privacy policies (e.g. information overload, extreme length of notices, impenetrable walls of texts, etc.) and consent requests (e.g. excessive quantity, lack of meaningful choice, etc.) examined in Section 2.2. As such, they mostly can be regarded as design patterns that foster transparency, as they offer operational ways to translate GDPR's user-centered transparency requirements [30] into practical solutions. Thus, these patterns represent the inform strategy, while the patterns about consent requests pertain to the control strategy [154]: the two strategies are connected, as informing data subjects about data collection should in principle allow them to better control such practice [69]. The legal design patterns presented in the following section constitute a potential bridge between the legal domain and the information and interaction design domain, with relevance for the engineering domain as well.

5.2.3 HCI Privacy Design Patterns

Whilst purely technical solutions have been investigated at length, design patterns exploring innovative ways to implement the inform and control strategies lack a considerable number of examples. The PrimeLife project produced and published a set of human-computer interaction (HCI) patterns [108], that contain some relevant examples and that, in some cases, have also been tested with users. Indeed, a peculiarity of this pattern collection is the attention devoted to the usability of such patterns for Privacy Enhancing Technologies. The design patterns presented in the following partially overlap with these HCI patterns, especially for what concerns the naming

Only the following patterns are published in the library (last access 29 June 2018): privacy icons, multi-layered notices, structured layout, and visual interface for active consent. These patterns have been however redefined, specified and modified in the present contribution, thus constitute original work. A first, rudimentary version of the patterns described in the following section was presented at the International Legal Informatics Symposium IRIS 2018, in Salzburg. A refined version of such patterns has been published in [252].

provided to the pattern categories related to the visualization of privacy information (e.g. privacy icons), to privacy policies (e.g. the privacy policy display pattern that corresponds to the multi-layered format), and to interaction (e.g. informed consent). Notwithstanding the terminological resemblance, the classification proposed here is different: the language patterns (Sec. 5.2.6) are primarily based on humans' verbal skills, the visualization patterns (Sec. 5.2.7) are based on the preponderance of visual non-linguistic elements, while the interaction patterns (Sec. 5.2.8) concern the necessity of user's interaction with the system.

However, in the following, the HCI privacy patterns of [108] that bear some resemblance to the scope of our research are elaborated, updated to the legal framework offered by the GDPR, integrated with authoritative sources from the Article 29 WP, and further specified. At the same time, some assumptions are explicitly rejected: for example, there are two patterns related to privacy icons, one for privacy icons as autonomous elements for any kind of application and an additional pattern for icons meant to appear on privacy policies. Firstly, our privacy icons pattern reunites the two perspectives and offers one unique privacy icon set with both functions. In the second place, we contest the presumed self-explicable nature of the graphical symbols, as also the same PrimeLife's researchers have later acknowledged (*see* [125]): not because self-explicability is not a desirable characteristic, but rather because it is concretely unattainable (*see* Section 6.1). For what concerns the informed consent pattern, it has been updated to meet the GDPR's newly introduced requirements and parceled out between two different patterns: active choice and specific consent. The privacy aware wording pattern has been improved in the transparent language pattern with more specific indications about clear and plain language.

5.2.4 Legal Design Patterns for Transparency and Consent

The translation of privacy requirements into applicable solution starts with an exploration of the space of design solutions [58]: the analysis presented in the following is based on an examination of the existing landscape. The exploration is focused on those solutions that have a clear human-centered focus, i.e. those interventions that attempt to inform individuals and put them in control as opposed to merely fulfil legal requirements. The analysis started in [134] and continuing in these pages is limited to the online environment, specifically to websites. Thus, privacy communication in software programs, apps and IoT devices is not analyzed, although many of these devices contain a link to an online privacy policy on a website. Moreover, pure academic research is also not included (for instance comics for privacy notices [178]): exclusively in-use solutions have been surveyed. Finally, the patterns proposed in the next few pages have a clear European (i.e. GDPR) focus.

Thus, the categorization proposed in the next pages is based on a definition of patterns as "[U]seful techniques in terms of the functional problem they aim to solve" [291, p.20]. The problems with privacy-related communication and consent that were identified in the literature review of the Chapter 2 are here matched to the different existing patterns that aim to solve them (*see* Fig. 5.1). As it will become evident, there is no one-to-one correspondence. Rather, each of the proposed strategies can contribute to solve at least one problem, but even more. Conversely, the solution to a problem can be attained by more than one pattern or by a combination of them. Table 5.1 summarizes the problems identified in the analysis reported in the first chapter and proposes high-level solutions, that will be declined in patterns in the following sections. Some problems concern the privacy policies at the individual level: non-readership, language, absent layout, the length of text, the timing of presentation, but also the content of the notices in terms of complexity and familiarity on the topics. Other problems consider the notices taken in their totality, such as the notice fatigue derived from the onslaught of notices. A last set of problems consider consent, with individual issues concerning the take-it-or-leave-it-approach and all-encompassing consent forms, whereas an additional problem concerns the quantity of consents. Furthermore, a classification of patterns is derived by three dimensions: language, visualization, and interaction. This can be seen as a move toward multi-sensory law [120]: the digital age offers rich possibilities to explore the law not only in verbal, but also in visual, auditory and tactile formats.

Problem	**Possible solution**
Non-readership: Users lack the motivation to read privacy notices, usually as a consequence of a variety of other issues outlined below, for instance the fact that all privacy policies look the same	Attract reader's attention
Language complexity: The language of privacy communication can be legalistic and unnecessary complex	Offer information in an intelligible manner; visually suggest or exemplify the meaning of the terms
Lack of audience-tailoring: Privacy communication is not designed with a specific user in mind, it is rather written by lawyers for lawyers	Offer meaningful information to the specific user
Vagueness of terms: Privacy terms can be open to multiple interpretations and leave the reader puzzled about their intended meaning	Clearly and unambiguously indicate if a certain data practice will happen or not

Wall of text: Privacy policies can be displayed as impenetrable texts, without any information architecture (e.g. paragraphs, headlines), thus hindering navigation and information-finding	Improve ease of navigation and skimmability
Excessive length: The text of privacy policies can be very long and cause information overload	Avoid information fatigue
Wrong timing: Privacy notices are presented at the time of data collection, causing hindrance to the primary task or too much distance in time to inform the privacy decision	Avoid nuisance factor
Lack of familiarity: Individuals lack necessary experience and knowledge to understand and assess privacy-related information and their consequent privacy decisions	Make the user more knowledgeable
Processing complexity: In the era of big data, it is hard to determine and explain how and why personal data is processed	Manage the number and complexity of big data practices in an intelligible manner
Notice fatigue: Individuals receive an onslaught of notices, which causes habituation effects	Manage the enormous number of privacy notices

Difficult comparability: There is no standard manner to present information across different privacy policies	Enhance information finding and comparability across services
Lack of meaningful choice: Certain services adopt a take-it-or-leave-it-approach	Do not bound users to give up unessential data to use the service
All-encompassing consent: Consent is asked globally	Provide the choice to give consent for single, specific processing purposes
Consent fatigue: Individuals receive an onslaught of consent requests, which causes habituation effects	Direct user's attention to the consent request

Table 5.1: Problems identified in the literature review in Section 2.2, their definition and corresponding high-level solution

5.2.5 The Pattern Structure

In the following, it is proposed an unpublished classification of design patterns (certain patterns do not serve one single function, but rather multiple ones). Each pattern has a structure that was adapted from the privacy patterns' online repository `https://privacypatterns.org/` and integrated with other useful information for their application. Each pattern presents the following structure:

Summary: defines the pattern;

Problem: lists the existing problem(s) (*see* Fig. 5.1) that the pattern aims to solve;

Solution: describes how the pattern can solve the problem;

Goals: lists the goal(s) the pattern is intended to achieve, in response to the problem(s);

Constraints and consequences: describe restrictions and aspects to which attention should be particularly devoted;

Modality: concerns the manner how the information is provided according to [260]: visual, auditory, machine-readable;

Legal reference: concerns whether the approach is suggested in a regulation or in an official opinion issued by supervisory authorities, in a purely European perspective.

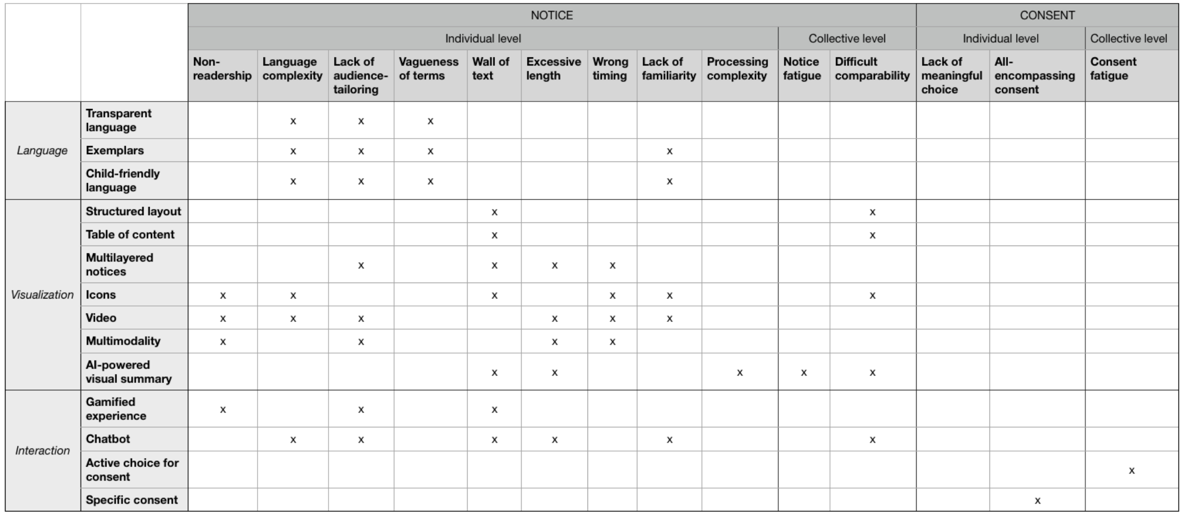
Each pattern can in principle solve one or even more problems: the following paragraphs attempt to provide a classification of the existing landscape, but a part from those patterns that are already widely adopted or that have undergone experimentation in similar domains, such as contracts, it is not yet demonstrated that these patterns actually contribute to solve the problem. The proposed classification can be therefore subject to change, as more evidence on the actual use and efficacy of the patterns is gathered.

5.2.6 Language Patterns

5.2.6.1 Transparent Language

Summary: Use clear and plain language that make the information easily comprehensible to anyone, especially laypeople.

Problem: Language complexity; lack of tailoring to the intended audience; vagueness of terms.



		NOTICE											CONSENT		
		Individual level									Collective level		Individual level		Collective level
		Non-readership	Language complexity	Lack of audience-tailoring	Vagueness of terms	Wall of text	Excessive length	Wrong timing	Lack of familiarity	Processing complexity	Notice fatigue	Difficult comparability	Lack of meaningful choice	All-encompassing consent	Consent fatigue
Language	Transparent language		x	x	x										
	Exemplars		x	x	x				x						
	Child-friendly language		x	x	x				x						
Visualization	Structured layout					x						x			
	Table of content					x						x			
	Multilayered notices			x		x	x	x							
	Icons	x	x			x		x	x			x			
	Video	x	x	x			x	x	x						
	Multimodality	x		x			x	x							
	AI-powered visual summary					x	x			x	x	x			
Interaction	Gamified experience	x		x		x									
	Chatbot		x	x		x	x		x			x			
	Active choice for consent														x
	Specific consent													x	

Figure 5.1: Classification of existing or emerging patterns according to the problem they can solve

Solution: Provide the information in the simplest possible language and in short sentences. Avoid complex structures and jargon terms. Check readability scores. Use concrete, direct lexical and syntactic structures, as opposed to vague and ambivalent terms.

Goals: Offer information in an intelligible manner; offer meaningful information to the specific user; clearly and unambiguously indicate if a certain data practice will happen or not.

Constraints and consequences: Clear and plain language does not necessarily mean trivial and oversimplified language. However, much effort must be devoted to finding the right wording with an acceptable trade-off between comprehensibility and preciseness. Some technical or legal specific terms (e.g. encryption) have no synonyms. In this case, provide the term and an explanation in simple words. As a result, the communication might appear even longer than before the simplification.

Modality: Visual

Legal reference: Recital 42 and Article 12.1 GDPR; Article 29 WP [30].

An example of clear and plain language that also contains an intelligible explanation of a legal-technical term is: "You create typeforms, and get data in the form of answers from respondents. You also handle data, then, making you a 'data controller'."[6]. As for what concerns concrete and direct language: "[W]e never store your credit card information and it never touches our server even for a few milliseconds"[7].

[6] Original example from Typeform's privacy policy https://admin.typeform.com/to/dwk6gt (Accessed on June 29, 2018).

[7] Original example from the Interaction Design Foundation's privacy policy https://www.interaction-design.org/about/privacy?utm_source=newsletter&utm_medium=email&utm_content=letter2018-05-24&utm_campaign=all (Accessed on June 29, 2018)

5.2.6.2 Exemplars

Summary: Provide an understandable example to clarify legal-technical terms or to make abstract concepts more tangible.

Problem: Language complexity; lack of tailoring to the intended audience; vagueness of terms; lack of familiarity.

Solution: Illustrative examples help users to transform abstract or unfamiliar legal information into more concrete experiences, that they might have lived or that they can easily imagine. They also provide practical instances of general categories and can be tailored to the intended audience of a certain service.

Goals: offer information in an intelligible manner; offer meaningful information to the specific user; explain in exact terms what abstract or complex notions mean in practice; make the user more knowledgeable (e.g. about data practices).

Constraints and consequences: The choice of examples must be relevant to the intended audience. Socio-demographics characteristics of the service's users (e.g. age) or the context (e.g. the type of service provided) can be leveraged, while the example's efficacy should be tested with the intended audience itself. Naming one example of a class must not hide other practices (framing effect): for instance, it is unlawful to exclusively mention privacy-friendly privacy practices, while omitting the risky ones.

Modality: Visual

Legal reference: /

A clear exemplar, tailored to the audience of the service and illustrating the data practices is the following: “The personal data and any information you provide us with [...] will be used, among other purposes, to: [...] send transactional email, e.g. send you an email when you have successfully earned a Course Certificate with a copy of your Certificate"[8].

5.2.6.3 Child-friendly Language

Summary: Design a child-tailored privacy policy if you offer services to under-aged users.

Problem: Language complexity; lack of tailoring to the intended audience; vagueness of terms; lack of familiarity.

Solution: The language should be addressed specifically to children and teenagers and take into consideration their level of cognitive development, which is different from that of an adult. Moreover, they should be made aware of their digital rights, while the consequences of their online choices in and the basics of privacy and data protection should be explained. Provide examples that are meaningful for their age. Optionally: provide child-tailored presentation modalities (e.g. comic strips).

Goals: offer information in an intelligible manner; offer meaningful information to the specific user; clearly and unambiguously indicate if a certain data practice will happen or not; make the user more knowledgeable (e.g. about data practices).

Constraints and consequences: Child-friendly language does not necessarily mean trivial and oversimplified language. However, much effort must be devoted to finding the right wording with an acceptable trade-off between comprehensibility and preciseness/correctness for a child or teenager. Some technical or legal specific terms (e.g. encryption) have no synonyms. In this case, provide the term and an explanation in simple words, that is relevant and meaningful for children and teenagers. As a result, the communication might appear even longer than before the simplification.

[8] From the Interaction Design Foundation's privacy policy, *see* footnote above (Accessed on June 29, 2018).

Modality: Visual.

Legal reference: Recital 38 and 59 Article 12.1 GDPR; Article 29 WP [30].

The Instagram's terms of service have been redrafted with attention to child-friendly language [70, p.10], e.g. "don't bully anyone or post anything horrible about people", on the model of the UN Convention on the Rights of the Child in child-friendly Language [283].

5.2.7 Visualization Patterns

5.2.7.1 Structured Layout

Summary: Organize the privacy policy in a consistent layout, where each topic is covered in a specific, labelled section[9].

Problem: Wall of text; difficult comparability.

Solution: Divide the privacy policy in thematic sections and assign a meaningful heading to each section. Typically, a privacy policy describes what, how and why personal data are used, and where and for how long they are processed. The information items that must be disclosed according to the GDPR are listed in Articles 13-14. Ideally, mark-up editors can be used to add machine-readable meanings to each section.

Goals: improve ease of navigation and skimmability; enhance information finding and comparability.

Constraints and consequences: The content must be rearranged according to the topics and an illustrative heading must be found.

Modality: Visual, machine-readable (optional).

[9] In its Guidelines to Transparency [30], the Article 29 WP collapses under the name "layered privacy statement/notices" a few strategies that are here proposed separately and specifically: structured layout, navigable table of content, and multi-layered notices. It was deemed necessary to name and describe them individually because they attempt to achieve overlapping but different goals.

Legal reference: Articles 13-14 GDPR

5.2.7.2 Navigable Table of Content

Summary: Place a navigable menu at the beginning of the page, where each item offers quick navigation to the corresponding section in the privacy policy.

Problem: Wall of text; difficult comparability.

Solution: A table of content at the beginning of the privacy policy that lists the headings of its different sections, thus offering an overview of the topics covered. It also supports the navigability of the document if each item has an hyperlink that connects it to the relevant section in the text. This structure can be particularly profitable for an efficient display on small screens.

Goals: Improve ease of navigation and skimmability; enhance information finding and comparability.

Constraints and consequences: The sections' headings must be carefully chosen in order to be meaningful and allow easy navigation.

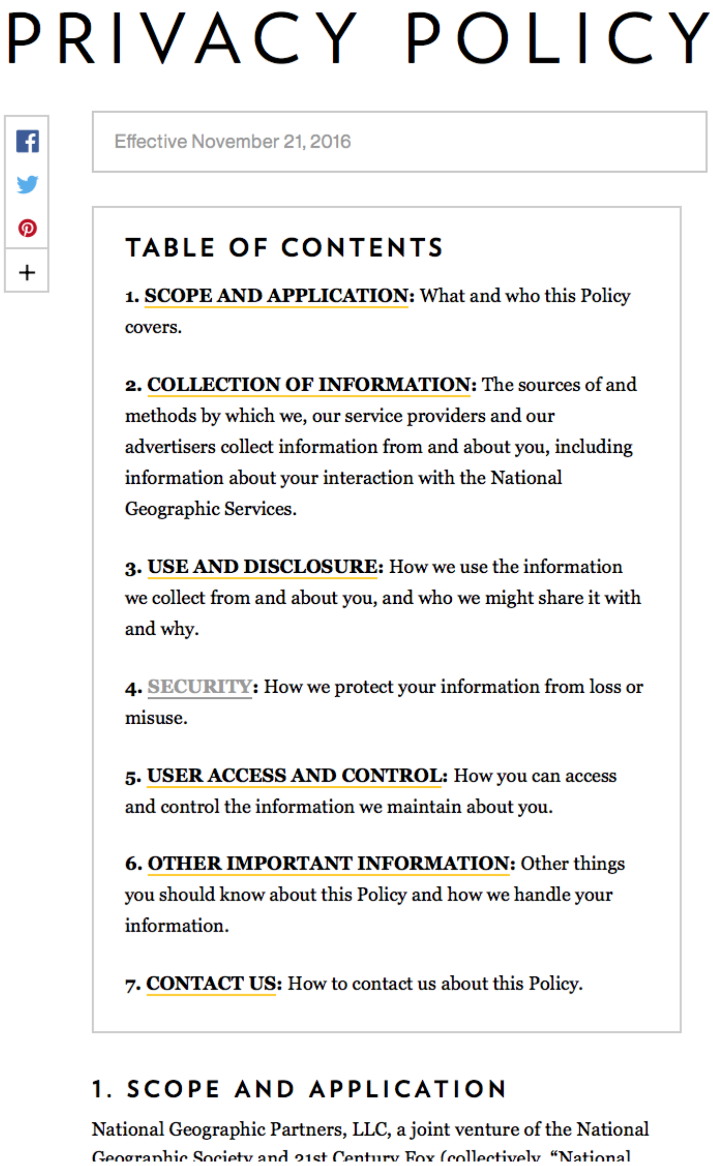
Modality: Visual, machine-readable.

Legal reference: /

5.2.7.3 Multi-layered Notices

Summary: The information is distributed on different layers, where the first layer offers an overview of the privacy policy, while more details are contained in the additional layers, that can also be explored.

Problem: Lack of audience-tailoring; wall of text; excessive length; wrong timing.



PRIVACY POLICY

Effective November 21, 2016

TABLE OF CONTENTS

1. **SCOPE AND APPLICATION:** What and who this Policy covers.

2. **COLLECTION OF INFORMATION:** The sources of and methods by which we, our service providers and our advertisers collect information from and about you, including information about your interaction with the National Geographic Services.

3. **USE AND DISCLOSURE:** How we use the information we collect from and about you, and who we might share it with and why.

4. **SECURITY:** How we protect your information from loss or misuse.

5. **USER ACCESS AND CONTROL:** How you can access and control the information we maintain about you.

6. **OTHER IMPORTANT INFORMATION:** Other things you should know about this Policy and how we handle your information.

7. **CONTACT US:** How to contact us about this Policy.

1. SCOPE AND APPLICATION

National Geographic Partners, LLC, a joint venture of the National Geographic Society and 21st Century Fox (collectively, "National

Figure 5.2: Example of navigable table of content. Source: National Geographic

Solution: Adopt a multi-layered approach to disclosures: instead of providing the entirety of information on one page, identify the most relevant, essential items and insert them on the first layer, while leaving more details and explanations to explorable layers on demand. Following recital 39 GDPR, the Article 29 WP suggests to include details about the purposes of processing, the identity of the controller, and a description of data subjects' rights. It is also possible to think of the different layers as addressing different audiences: the first layer for those data subjects that desire an overview, while the second layer is addressed to supervisory authorities, or likewise, and to those data subjects that desire the extended explanation. Timing is also a relevant dimension in this context: the first layer can be shown when the user is executing a different task, while the second layer can be provided on demand.

Goals: Offer meaningful information to the specific user; improve ease of navigation and skimmability; avoid information fatigue; avoid nuisance factor.

Constraints and consequences: The information provided on the different layers must be on its whole consistent and harmonized, i.e. the information in one layer cannot conflict with the information on a different layer. The first layer must not include only fair terms, whereas unfair or risky practices are buried down into the other layers. This multi-layered structure offers compliance in its totality.

Modality: Visual, machine-readable (optional).

Legal reference: Recital 39 GDPR; Article 29 WP [22, 30].

5.2.7.4 Progress mechanism

Summary: Display a mechanism showing the progress of the user through the privacy policy.

Problem: Non-readership.

Solution: The advancement of a progress bar or a similar mechanism e.g. showing a percentage displays the proportional amount of work that the user has completed, i.e. with respect to the fulfilment of the task of reading a sheer amount of privacy information. Privacy information can be organized in different chunks and in different windows, but it is important to provide orientation to the user, i.e. suggest her what has been already accomplished and the estimated time of reading the rest of the information.

Goals: Provide a tangible manner to show users their progress and, thereby, support their motivation to read.

Constraints and consequences: Although displaying a progress mechanism is always meaningful, it cannot be expected that users will deterministically read the whole privacy policy.

Modality: Visual.

Legal reference: /

An example of progress bar is showed in Fig. 5.8.

5.2.7.5 Icons

Summary: Icons accompany the verbal privacy policy and visually suggest where a specific piece of information in the long text can be found.

Problem: Non-readership, language complexity, wall of text, wrong timing, lack of familiarity, difficult comparability.

Solution: Include icons in the privacy policy or use them in combination with text on a layered notice to provide a quick overview of the data processing.

Goals: Attract reader's attention; visually suggest or exemplify the meaning of the terms; improve ease of navigation and skimmability; avoid nuisance factor; enhance information finding and comparability.

Constraints and consequences: icons should always be accompanied by a textual explanation, since they are generally not self-explanatory and users might be unfamiliar with them; this risk lowers if the icons pertain to a shared, standard visual vocabulary.

Modality: Visual, machine-readable.

Legal reference: Art 12.7 GDPR; Article 29 WP [30].

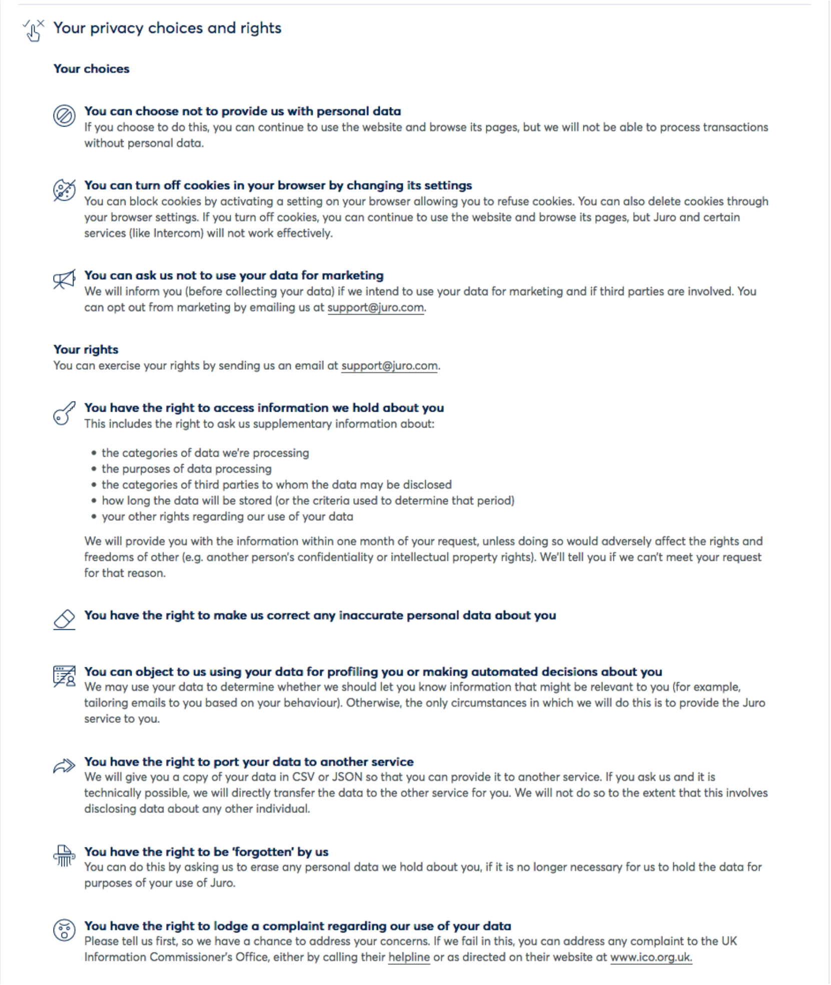
Juro offers a great example of a privacy policy that has been redesigned by a multi-disciplinary team, among which Stefania Passera, that has applied the legal design principles from best practice [192] to a traditional online privacy policy. Juro's privacy policy integrates icons into good information architecture (*see* Fig. 5.3) and layered structure.

5.2.7.6 Videos

Summary: The main points of the privacy policy are communicated through a video.

Problem: Non-readership; language-complexity; lack of audience tailoring; excessive length; wrong timing; lack of familiarity.

Solution: Provide a short, introductory video that explains the main data practices of the organization on the privacy policy page and possibly on a video-sharing platforms like Youtube. The video, that can be animated or not, is conceived as a summary of the most fundamental aspects of the entity's data processing. As such, it seems suitable to be used in those occasions where an overview of the processing would prove useful, but a long privacy policy would be considered a nuisance factor. The video can catch the attention and also be perceived as a less time-consuming activity than reading the whole document, thus encouraging its viewing. Unlike other means, a video also has the capacity to convey the tone and feeling of the relationship of the organization with the data



Your privacy choices and rights

Your choices

You can choose not to provide us with personal data
If you choose to do this, you can continue to use the website and browse its pages, but we will not be able to process transactions without personal data.

You can turn off cookies in your browser by changing its settings
You can block cookies by activating a setting on your browser allowing you to refuse cookies. You can also delete cookies through your browser settings. If you turn off cookies, you can continue to use the website and browse its pages, but Juro and certain services (like Intercom) will not work effectively.

You can ask us not to use your data for marketing
We will inform you (before collecting your data) if we intend to use your data for marketing and if third parties are involved. You can opt out from marketing by emailing us at support@juro.com.

Your rights
You can exercise your rights by sending us an email at support@juro.com.

You have the right to access information we hold about you
This includes the right to ask us supplementary information about:

- the categories of data we're processing
- the purposes of data processing
- the categories of third parties to whom the data may be disclosed
- how long the data will be stored (or the criteria used to determine that period)
- your other rights regarding our use of your data

We will provide you with the information within one month of your request, unless doing so would adversely affect the rights and freedoms of other (e.g. another person's confidentiality or intellectual property rights). We'll tell you if we can't meet your request for that reason.

You have the right to make us correct any inaccurate personal data about you

You can object to us using your data for profiling you or making automated decisions about you
We may use your data to determine whether we should let you know information that might be relevant to you (for example, tailoring emails to you based on your behaviour). Otherwise, the only circumstances in which we will do this is to provide the Juro service to you.

You have the right to port your data to another service
We will give you a copy of your data in CSV or JSON so that you can provide it to another service. If you ask us and it is technically possible, we will directly transfer the data to the other service for you. We will not do so to the extent that this involves disclosing data about any other individual.

You have the right to be 'forgotten' by us
You can do this by asking us to erase any personal data we hold about you, if it is no longer necessary for us to hold the data for purposes of your use of Juro.

You have the right to lodge a complaint regarding our use of your data
Please tell us first, so we have a chance to address your concerns. If we fail in this, you can address any complaint to the UK Information Commissioner's Office, either by calling their helpline or as directed on their website at www.ico.org.uk.

Figure 5.3: The section about data subjects' rights of Juro's privacy policy [11].

subjects. Moreover, it can provide information for visually impaired individuals or individuals with low or in-existent levels of literacy.

Goals: Attract reader's attention; visually suggest or exemplify the meaning of the terms; offer meaningful information to the specific user; avoid information fatigue; avoid nuisance factor; make the user more knowledgeable (e.g. about data practices).

Constraints and consequences: The video can communicate the fundamental data practices of an organization, akin to the first layer of information in a multi-layered approach, and has therefore to be combined with a complete, written privacy policy for compliance. The video must not focus exclusively on fair terms, whereas unfair or risky practices are buried down into the written document, since individuals might watch the video but not read the document. Therefore, the choice about what to include and what to leave out must be carefully made. Also technical constraints should be considered.

Modality: Visual, auditory.

Legal reference: Article 29 WP [30].

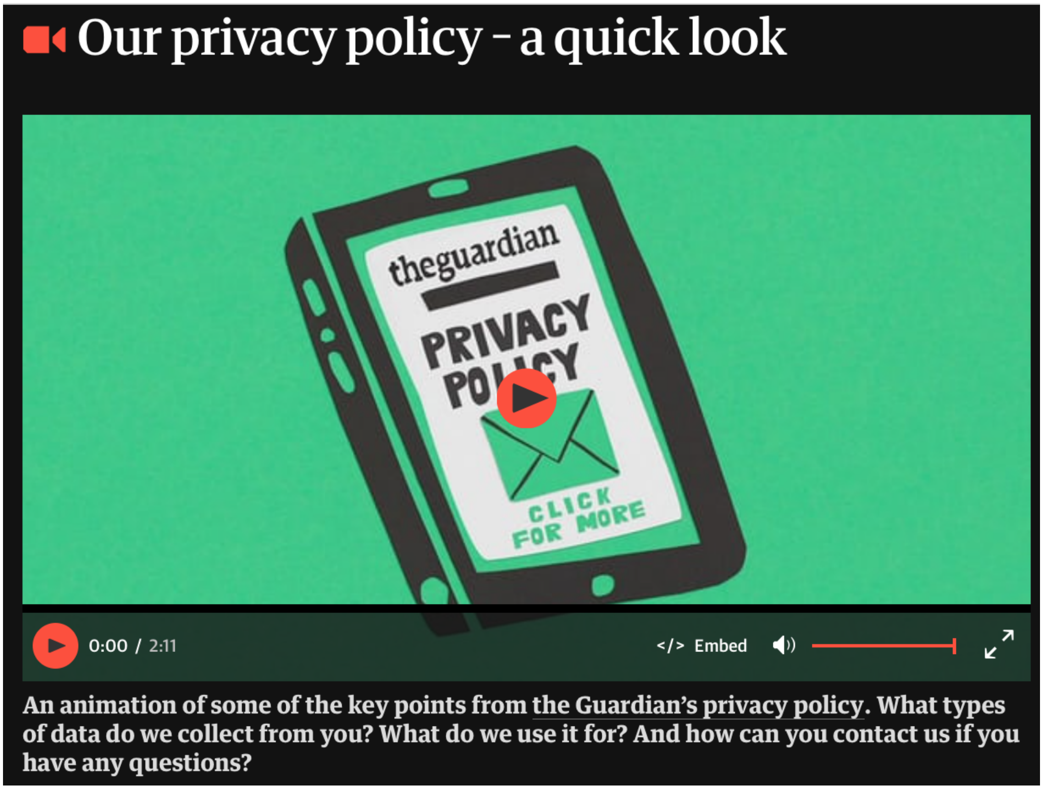
There exist a few examples of organizations that have expressed their privacy practices through a video, for instance Google, Linkedin, Easyjet [86] and The Guardian [128] (*see* Fig. 5.4).

5.2.7.7 Multimodality

Summary: Privacy information is conveyed through multiple channels, e.g. auditory and visual.

Problem: Non-readership; lack of audience tailoring; excessive length; wrong timing.

Solution: Offer multiple channels to provide of the information about privacy and data protection, for instance in written form (e.g. document)



Our privacy policy - a quick look

An animation of some of the key points from the Guardian's privacy policy. What types of data do we collect from you? What do we use it for? And how can you contact us if you have any questions?

Figure 5.4: The Guardian's video introducing its privacy policy's principles.

and visual form (e.g. video). Indeed, users might not want or might be unable to read the full, written privacy policy due to time constraints, device constraints, or even physical impairment. Other actors might well need the complete privacy policy, e.g. for scrutiny. According to their characteristics, to their goals, or to the context (e.g. task, device), users can thus choose their preferred channel to be informed.

Goals: Attract reader's attention; offer meaningful information to the specific user; avoid information fatigue; avoid nuisance factor.

Constraints and consequences: The information on the different channels must be carefully harmonised and it must be clear that different channels convey different aspects of the privacy information.

Modality: visual, auditory, machine-readable.

Legal reference: Article 29 WP [30].

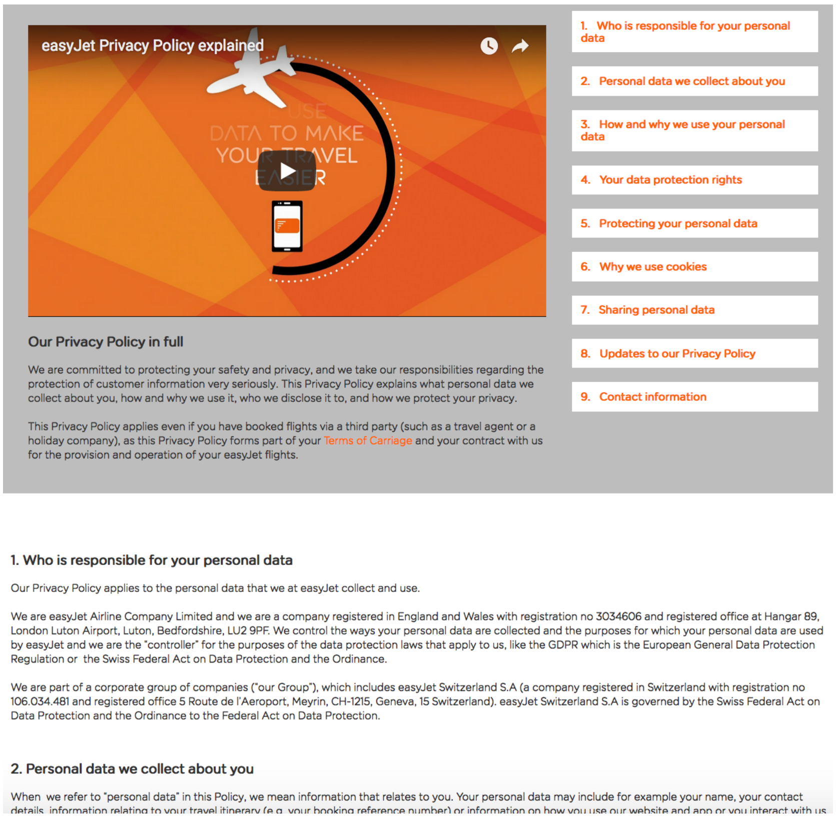
Easyjet [86] has adopted a blended approach to privacy communication and makes use of many patterns that have been hitherto described. Fig. 5.5 shows that the company prominently presents a video, while at its side there is a navigable table of content that is linked to the written privacy policy right below. Other services have decided to provide a traditional privacy policy, with a face-to-face translation into plain terms (*see* Fig. 5.6).

5.2.7.8 AI-powered Visual Summary

Summary: Let an AI analyze the practices described in a privacy policy and visually summarize them.

Problem: Wall of text; excessive length; processing complexity; notice fatigue, difficult comparability.

Solution: Certain forms of artificial intelligence, if appropriately created and trained, can be able to automatically analyze a textual privacy



1. Who is responsible for your personal data

2. Personal data we collect about you

3. How and why we use your personal data

4. Your data protection rights

5. Protecting your personal data

6. Why we use cookies

7. Sharing personal data

8. Updates to our Privacy Policy

9. Contact information

Our Privacy Policy in full

We are committed to protecting your safety and privacy, and we take our responsibilities regarding the protection of customer information very seriously. This Privacy Policy explains what personal data we collect about you, how and why we use it, who we disclose it to, and how we protect your privacy.

This Privacy Policy applies even if you have booked flights via a third party (such as a travel agent or a holiday company), as this Privacy Policy forms part of your Terms of Carriage and your contract with us for the provision and operation of your easyJet flights.

1. Who is responsible for your personal data

Our Privacy Policy applies to the personal data that we at easyJet collect and use.

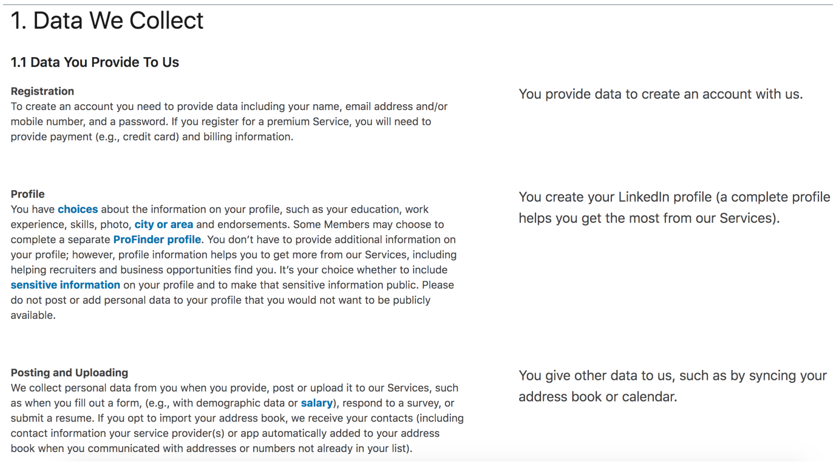
We are easyJet Airline Company Limited and we are a company registered in England and Wales with registration no 3034606 and registered office at Hangar 89, London Luton Airport, Luton, Bedfordshire, LU2 9PF. We control the ways your personal data are collected and the purposes for which your personal data are used by easyJet and we are the "controller" for the purposes of the data protection laws that apply to us, like the GDPR which is the European General Data Protection Regulation or the Swiss Federal Act on Data Protection and the Ordinance.

We are part of a corporate group of companies ("our Group"), which includes easyJet Switzerland S.A (a company registered in Switzerland with registration no 106.034.481 and registered office 5 Route de l'Aeroport, Meyrin, CH-1215, Geneva, 15 Switzerland). easyJet Switzerland S.A is governed by the Swiss Federal Act on Data Protection and the Ordinance to the Federal Act on Data Protection.

2. Personal data we collect about you

When we refer to "personal data" in this Policy, we mean information that relates to you. Your personal data may include for example your name, your contact details, information relating to your travel itinerary (e.g. your booking reference number) or information on how you use our website and app or you interact with us

Figure 5.5: Easyjet [86] offers two complementary ways to explore its data practices: in video and in written form.



1. Data We Collect

1.1 Data You Provide To Us

Registration To create an account you need to provide data including your name, email address and/or mobile number, and a password. If you register for a premium Service, you will need to provide payment (e.g., credit card) and billing information.	You provide data to create an account with us.
Profile You have **choices** about the information on your profile, such as your education, work experience, skills, photo, **city or area** and endorsements. Some Members may choose to complete a separate **ProFinder profile**. You don't have to provide additional information on your profile; however, profile information helps you to get more from our Services, including helping recruiters and business opportunities find you. It's your choice whether to include **sensitive information** on your profile and to make that sensitive information public. Please do not post or add personal data to your profile that you would not want to be publicly available.	You create your LinkedIn profile (a complete profile helps you get the most from our Services).
Posting and Uploading We collect personal data from you when you provide, post or upload it to our Services, such as when you fill out a form, (e.g., with demographic data or **salary**), respond to a survey, or submit a resume. If you opt to import your address book, we receive your contacts (including contact information your service provider(s) or app automatically added to your address book when you communicated with addresses or numbers not already in your list).	You give other data to us, such as by syncing your address book or calendar.

Figure 5.6: Linkedin [186] provides a privacy notice written in traditional, legal terms, and a vis-a-vis translation into transparent language.

policy according to pre-defined privacy-related categories and, consequently, offer an upfront, visual summary according to such categories. The results of the analysis can be then more easily compared across different services.

Goals: improve ease of navigation and skimmability; avoid information fatigue; manage the number and complexity of big data practices in an intelligible manner; manage the enormous number of privacy notices; enhance information finding and comparability.

Constraints and consequences: Such AI must be developed and trained according to the categories considered relevant for the users. Individuals must be aware that the system selects the information according to its training. Moreover, any system that automatically analyses texts and extracts meaning can be prone to error. An appropriate manner of visual representation that intelligibly displays the analysis' results must also be carefully designed.

Modality: visual, machine-readable.

Legal reference: /

To the best of our knolwedge, there exists only one example of this pattern: an online interactive interface (*see* Fig. 5.7) that displays the results of the automated analysis of privacy policies carried out by Polesis [145] (*see* Section 3.3.2.1). Through the interface, the user can either select one service listed in the database, or ask the analysis of another service's data practices. The results are shown as streams of colors indicating the presence of a certain practice. The thickness of the streams signifies the preponderance of the practices enabling comparability, for instance in terms of quantity of personal data that are passed on to third parties with respect to a different service. The interface is interactive and allows for the visual exploration of data practices.

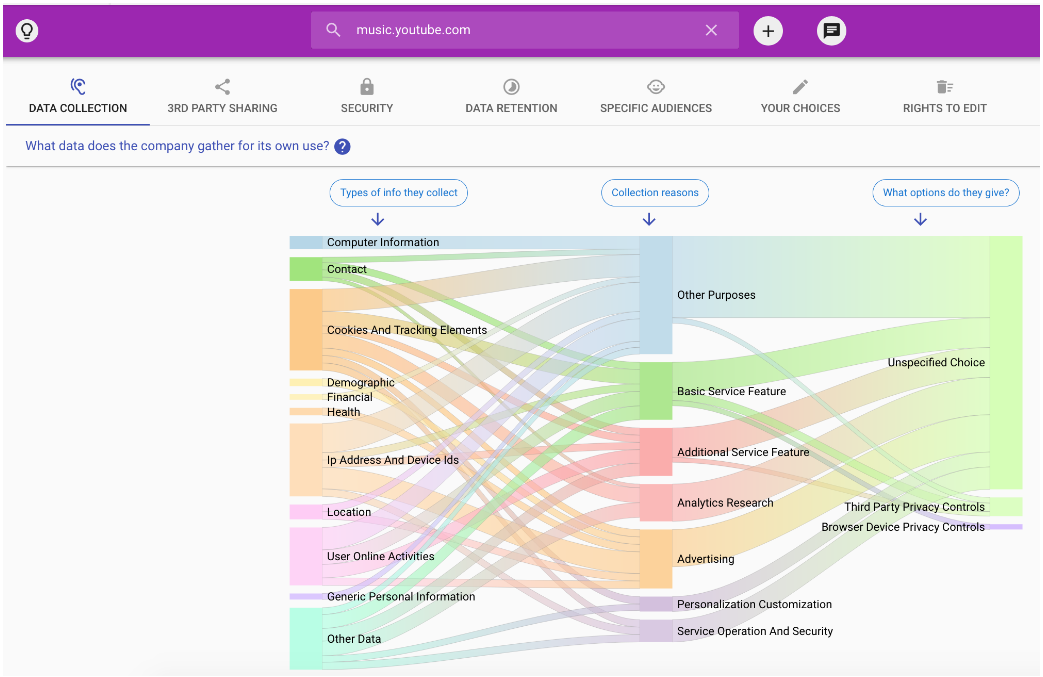

music.youtube.com
DATA COLLECTION
3RD PARTY SHARING
SECURITY
DATA RETENTION
SPECIFIC AUDIENCES
YOUR CHOICES
RIGHTS TO EDIT
What data does the company gather for its own use?
Types of info they collect
Collection reasons
What options do they give?
Computer Information
Contact
Cookies And Tracking Elements
Demographic
Financial
Health
Ip Address And Device Ids
Location
User Online Activities
Generic Personal Information
Other Data
Other Purposes
Basic Service Feature
Additional Service Feature
Analytics Research
Advertising
Personalization Customization
Service Operation And Security
Unspecified Choice
Third Party Privacy Controls
Browser Device Privacy Controls

Figure 5.7: The results of the automated analysis of privacy policies carried out by Polesis are displayed in an interactive, visual interface.

5.2.8 Interaction Patterns

5.2.8.1 Gamified Experience

Summary: Present the privacy practices in a gamified environment.

Problem: Lack of motivation to read; lack of audience-tailoring; wall of text.

Solution: Design an experience by making use of gamified mechanics that allows users gain a reward (e.g. in terms of points, badges, etc.) by exploring the privacy practices, thus enhancing their motivation to read.

Goals: Attract reader's attention; offer meaningful information to the specific user; improve ease of navigation and skimmability.

Constraints and consequences: The gamified exploration of the privacy practices must not be compulsory, but rather seen as an added value:

it must reflect the user's free choice to be informed, otherwise it risks to be considered as a nuisance. This is why it must always be accompanied by a traditional privacy policy, which is the authoritative version. The use of mechanics from gamification must be well integrated in the goals and philosophy of the service and offer a reward that is meaningful for the users, who must also be able to spend it somewhere. This is why probably only specific organizations can smoothly integrated such mechanisms into their service, e.g. videogames or services based on gamified elements.

Modality: visual, auditory (optional).

Legal reference: /

To the best of our knowledge, there exists only one concrete example of this pattern[10]: PrivacyVille, the gamified experience integrated into the Zynga's online videogames (*see* Fig. 5.8). In the city of PrivacyVille, each building represents a different data practice. The user is invited to explore every building and at the end, after a short comprehension quiz, she is awarded a certification that she can leverage in a company's videogame.

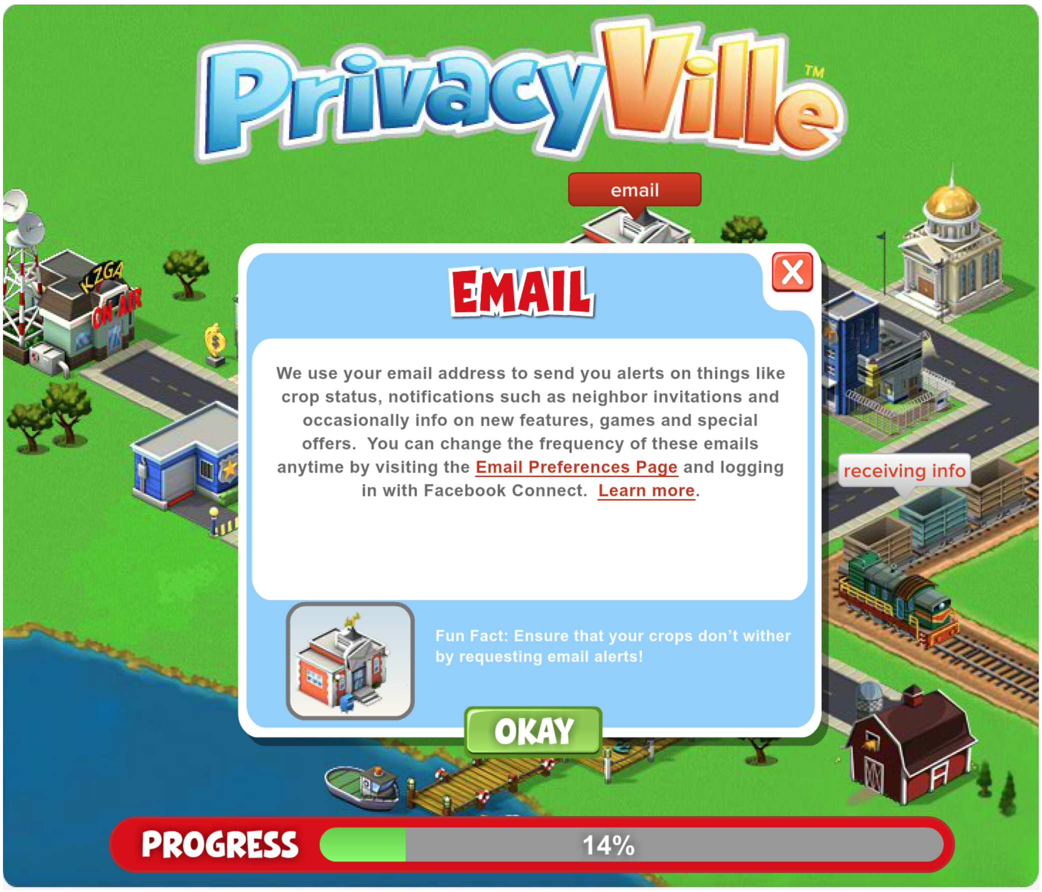
5.2.8.2 Question-Answering Chatbot

Summary: Exploration of a privacy policy in a personalized and interactive way through a conversation with a chatbot.

Problem: Language complexity; lack of audience-tailoring; wall of text; excessive length; lack of familiarity; difficult comparability.

Solution: A chatbot is a computer program that simulates human conversations through text chats or voice commands. If properly designed

[10] Unfortunately, at the time of writing this dissertation (July 2, 2018), the link to PrivacyVille [306] redirects the visitor to the general privacy policy and the gamified experience does not seem to be online anymore.



We use your email address to send you alerts on things like crop status, notifications such as neighbor invitations and occasionally info on new features, games and special offers. You can change the frequency of these emails anytime by visiting the Email Preferences Page and logging in with Facebook Connect. Learn more.

Fun Fact: Ensure that your crops don't wither by requesting email alerts!

Figure 5.8: A screenshot of PrivacyVille where the user is exploring the data practices linked to her e-mail. The progress-bar is an additional visual device that shows to the user her progression through the totality of terms that must be explored to earn the final certificate

and trained, it can answer users' questions about an organization's privacy practices in a reliable manner and in real-time. Thus, the user can find the information she is looking for easily and rapidly and the communication is tailored to her needs and interests. The chatbot can be offered by the organization itself or set up by a third party.

Goals: offer information in an intelligible manner; offer meaningful information to the specific user; improve ease of navigation and skimmability; make the user more knowledgeable (e.g. about data practices); enhance information finding and comparability.

Constraints and consequences: Depending on its level of sophistication, it shows different flexibility in understanding the questions and giving the answers. It is an intermediary between privacy policy text and user, thus the users should be made aware of the fact that the chatbot might not be completely certain about the answers and might provide an interpretation of the text.

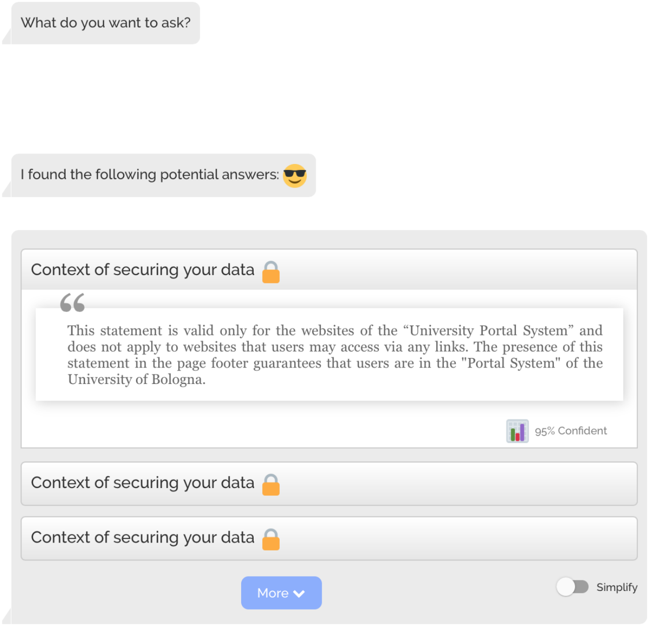
Modality: visual, auditory (optional), machine-readable.

Legal reference: /

To the best of our knowledge, Pribot [146, 145] is the only conversational agent explicitly dedicated to the data practices described in privacy policies. The chatbot only allows the user to select a question among a specific pool of pre-defined questions, according to the categories for which the classification algorythm has been trained. This specific chatbot makes use of an algorythm that analyzes the written information contained in a privacy policy, finds the relevant section where a topic is mentioned and proposes it to the user, showing a percentage that reveals its confidence and also employs emojis to mimic a conversational tone (*see* Fig. 5.9).

5.2.8.3 Active Choice

Summary: An interface with active opt-in for consent, instead of pre-ticked



What do you want to ask?

Do you use any form of encryption?

I found the following potential answers: 😎

Context of securing your data 🔒

This statement is valid only for the websites of the “University Portal System” and does not apply to websites that users may access via any links. The presence of this statement in the page footer guarantees that users are in the "Portal System" of the University of Bologna.

95% Confident

Context of securing your data 🔒

Context of securing your data 🔒

Figure 5.9: A screenshot of PriBot’s answers to the user’s question about a service’ privacy policy. As it is evident, the chatbot only provides the spans of text where the answer should be contained, leaving to the user the burden of finding and interpreting the answer

boxes.

Problem: Consent fatigue.

Solution: Design an interface with active choices for the user to unambiguously signify her consent to data collection and processing. Opt-in choices attract user's attention because she is actively mandated to take a decision and does not go unnoticed, contrary to the pre-ticked boxes that cause habituation and rely on the status quo bias.

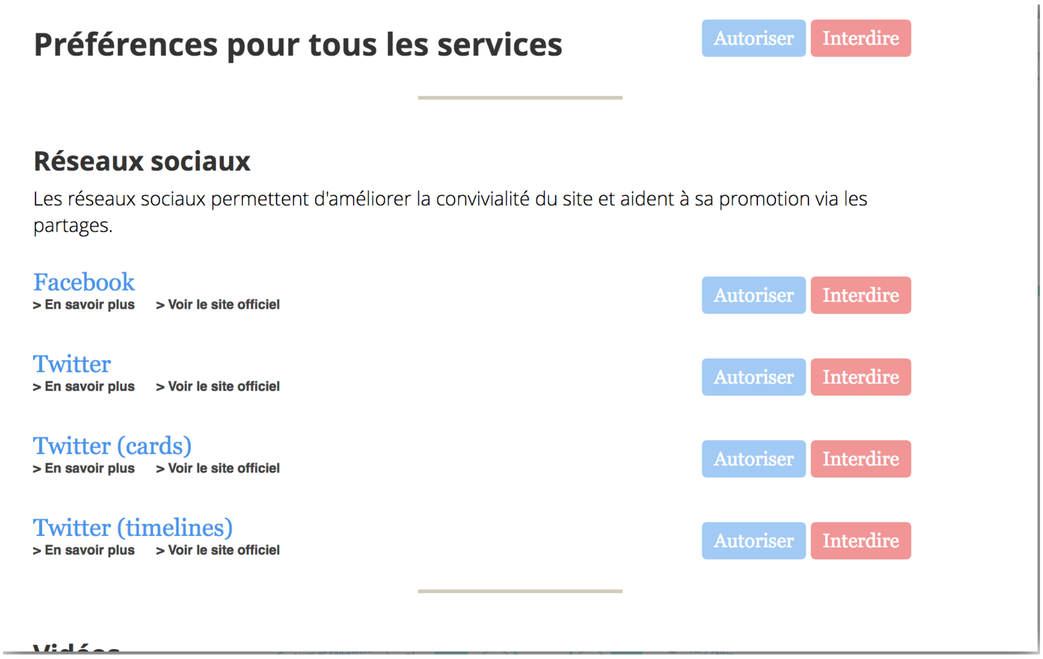
Goals: Direct user's attention to the consent request.

Constraints and consequences: the two answers (i.e. consent/don't consent, allow/don't allow, yes/no) should be given vertically so that individuals will assign the same weight to both of them, whereas when the choice is horizontal, individuals are more likely to choose the item on the right [160]. When the user receives too many consent requests or is forced to take a decision while carrying out a different task, she might live active choice as a nuisance factor. Good practice is therefore to offer consent management at a higher level, like the browser level.

Modality: visual.

Legal reference: Recital 32 and Art. 4 GDPR; Article 29 WP [26].

After the GDPR became applicable, a number of websites updated their cookie consent requests to offer a concrete, active choice instead of the widely adopted default opt-in. In particular, two forms of active consent have arisen: some entities offer an equal choice between accepting and refusing a certain data processing through two equally important buttons (*see* Fig. 5.10), whereas other entities have preferred to provide default opt-outs that can be easily turned into opt-ins at the individual's will (*see* Fig. 5.11).



Préférences pour tous les services

Réseaux sociaux

Les réseaux sociaux permettent d'améliorer la convivialité du site et aident à sa promotion via les partages.

Facebook
> En savoir plus > Voir le site officiel

Twitter
> En savoir plus > Voir le site officiel

Twitter (cards)
> En savoir plus > Voir le site officiel

Twitter (timelines)
> En savoir plus > Voir le site officiel

Vidéos

Figure 5.10: A screenshot of the active choice between consent authorization ('autoriser') and consent refusal ('intedire') about social media on the CNIL's website.

5.2.8.4 Specific Consent

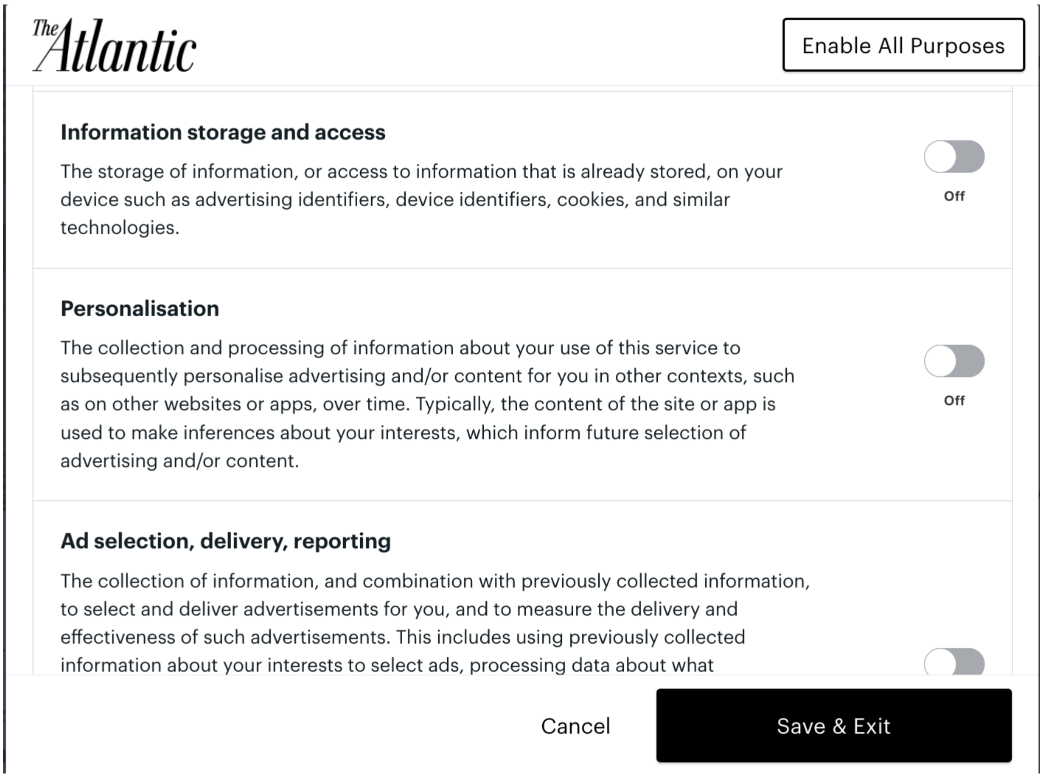
Summary: An interface for consent where each processing purpose is separated from the other

Problem: All-encompassing consent.

Solution: Design an interface for consent requests, where each processing purpose for which consent is required is displayed and explained prominently. For each processing purpose, provide an active choice to the user.

Goals: Provide the choice to give consent for single, specific processing purposes.

Constraints and consequences: If there are many different purposes, the user is forced to make a choice for each one of them, causing fatigue.



Information storage and access

The storage of information, or access to information that is already stored, on your device such as advertising identifiers, device identifiers, cookies, and similar technologies.

Off

Personalisation

The collection and processing of information about your use of this service to subsequently personalise advertising and/or content for you in other contexts, such as on other websites or apps, over time. Typically, the content of the site or app is used to make inferences about your interests, which inform future selection of advertising and/or content.

Off

Ad selection, delivery, reporting

The collection of information, and combination with previously collected information, to select and deliver advertisements for you, and to measure the delivery and effectiveness of such advertisements. This includes using previously collected information about your interests to select ads, processing data about what

Figure 5.11: A screenshot of the default opt-outs that the user must actively change into opt-ins to signify her consent to specific data processing activities on The Atlantic's website.

> Thus, it is good practice to provide an all-encompassing choice (e.g. enable all purposes, *see* Fig. 5.10) that the user can select if she does not want to decide for each.

Modality: visual.

Legal reference: Recital 32 and Art. 4 GDPR; Article 29 WP [26].

Website's consent requests usually concern cookies and have traditionally taken the shape, of a fictitious consent request where the user is forced to either accept all the processing purposes or go elsewhere. The GDPR has fostered innovation in this sense and many newly arisen solutions offer granular controls (*see* e.g. Fig. 5.11).

The classified efforts presented in these last pages exemplify that there exist attempts to innovate legal and privacy-related communication, as growing interest towards the usability and information design of legal documents testifies. However, apart from a handful of cases, many of the above described design patterns for privacy and data protection present rare and even unique examples, e.g. the gamified experience or the privacy chatbot, whilst consent is being implemented in several different ways. The great majority of leading websites employ exclusively the simplest strategies [134], such as structured layout, table of contents and summary tables, if any. It follows that much more focus on innovative practices is needed, whilst this analysis should be repeated after the GDPR came into effect to check whether the transparency obligation has fostered innovation on a large scale. This data, however, suggests that it would be inexact to claim that legal design patterns for privacy exist: there rather exist isolated solutions that have been adopted by a few single entities. This is why it is more exact to describe these patterns as candidate patterns: only their adoption and widespread application will reveal if they will actually become replicated patterns.

Future work will include the assessment about the effective adoption and the efficacy of such patterns, for instance in terms of usability assessments and measurable improvements (*see* also [69]). The evaluative dimension is critical also to demonstrate compliance with the GDPR's provisions on transparency requirements (*see* [30]): although several solutions might implement the same design pattern, it is not necessarily true that all will be equally effective at solving the problem - however, a safe threshold could be set as minimal requirement. For the icons pattern explored in the next chapter, it is hereby provided an experiment of evaluation and indications about further assessment. In addition, best practice shows that a more rigorous classification showing relationships among these patterns, and between these patterns and other existing ones is needed (e.g. following the types of pattern relationships presented in [58] and the best practices in [293]).

5.3 Dark Patterns

As demonstrated in the last section, technological and design interventions can make privacy-related information and options easier to understand and to use. Nevertheless, bad information design has deliberately, or simply out of ignorance, obscured information around privacy and data protection in unintelligible and hard-to-navigate documents such as privacy policies, while specific interface design choices such as default choices have nudged users to disclose information or to give consent to certain processing activities on their data. Thus, choice architecture can influence users towards desirable privacy goals, but it can also be exploited to nudge users towards less desired outcomes [15]. "Although most individuals are probably unaware of the diverse influences on their concern about privacy, entities whose interests depend on information revelation by others are not" cautions Acquisti [16]. Manipulation of subtle elements can influence users towards more sharing or make it harder to choose a privacy-friendly behavior and the few existing studies on the topic that will be introduced below seem to confirm this assumption.

The success of design patterns to record working solutions have also given rise to a research of the opposite sense: anti-patterns and dark patterns. Anti-patterns are solutions that should be avoided because represent bad practices and have negative consequences, while dark patterns are "malicious patterns that intentionally weaken or exploit the privacy of users, often by making them disclose personal data or consent against their real interest" [48, p. 237]. Hence, the first are the result of bad design choices that inadvertently trick the user, while the latter are patterns that purposedly lure the user into privacy-unfriendly behaviors.

Common practice for service providers is represented by the obscure strategy, that make it "hard or even impossible for data subjects to learn how their personal data is collected, stored, and processed". As it is evident, a privacy policy characterized by low levels of transparency realizes this strategy in an optimal manner. Another dark pattern that belongs to this strategy is bad defaults, i.e. default options at the moment of account creation that ease the sharing of personal data.

5.3.1 Common Examples

One month after the application of the GDPR, the Norwegian Consumer Council released a report on the dark patterns employed by Google, Facebook, and Microsoft [9]. Dark patterns are defined as "exploitative design choices" (p. 4) and are considered problematic from an ethical point of view "because they mislead users into making choices that are not in their interest, and deprive them of their agency" (p. 7). The patterns identified in the report are classified following [15], highlighting how each category of nudges can serve as a positive, lawful, ethically-oriented nudge towards the users' best interests, but also in the opposite sense.

The study found that the companies used a variety of techniques to purposedly deceive their users and lure them into more privacy-intrusive options, some of them explicitly prohibited by the GDPR. For example, Google and Facebook provide privacy-invasive defaults and make it harder for users to discover and choose privacy-friendly options, for instance by entailing these in actions that require more clicks and more time. Such options are explicitly tackled by the GDPR's principle of data protection by default (*see* Section 5.1). As for what concerns framing, some risks were downplayed by specific wordings, whereas privacy-preserving options were discouraged by making this choice seem ethically questionable or insecure. For example, Facebook framed the option of not activating facial recognition (i.e. a biometric data, hence sensitive data) as a risk for user's security: "if you keep face recognition turned off, we won't be able to use this technology if a stranger uses your photo to impersonate you" or as an unscrupulous choice: "[i]f someone uses a screen reader, they won't be told when you're in a photo unless you're tagged". Moreover, these services pressured their users to make privacy decisions at a specific time, making it impossible to postpone such decisions. Finally, Google and Facebook menaced their users with loss of functionality or account deletion, if the privacy updates were not accepted, which opposes the notion of freely given consent.

5.3.2 Cookies

As it is evident, choice architecture can nudge data subjects towards more disclosure, if used for unethical goals. Similar tactics are actively exploited by the marketing industry, for instance for what concerns default settings for cookies on websites. One striking example is provided by the cookie solution adopted by any Oath owned website (such as Yahoo, Tumblr, and Techcrunch), that will be described and critically analyzed in the following.

1. After the GDPR's application deadline, Europeans found a cookie wall on any Oath's website, reproduced in Fig. 5.12, that employs a number of strategies used to lure users into accepting the conditions. Firstly, users are explicitly instructed to "select Accept" after they have inspected the practices. The indication is placed prominently at the beginning of the page, whereas the possibility to manage cookies is buried down in text. It is also questionable if the length of the text acts as a malicious nudge, i.e. to discourage the user from reading the options and managing the settings. Framing is another employed technique: the headline stating "How data brings you better experiences" prominently describes the positive sides of accepting the cookies and disclosing data, while the company skates over other consequences and risks. The colors and position of the buttons at the end of the page also nudge the user toward accepting the practices *in toto*: the "OK" button is colored in a vivid light blue, meaning that it is active, and it is placed on the right-hand side, which is the usual position for buttons that make the user proceed to the next page. On the contrary, the "manage options" button is in a light grey and placed on the left, making it less visible and less relevant for the user. In other words, the user is indirectly nudged towards accepting the conditions and go

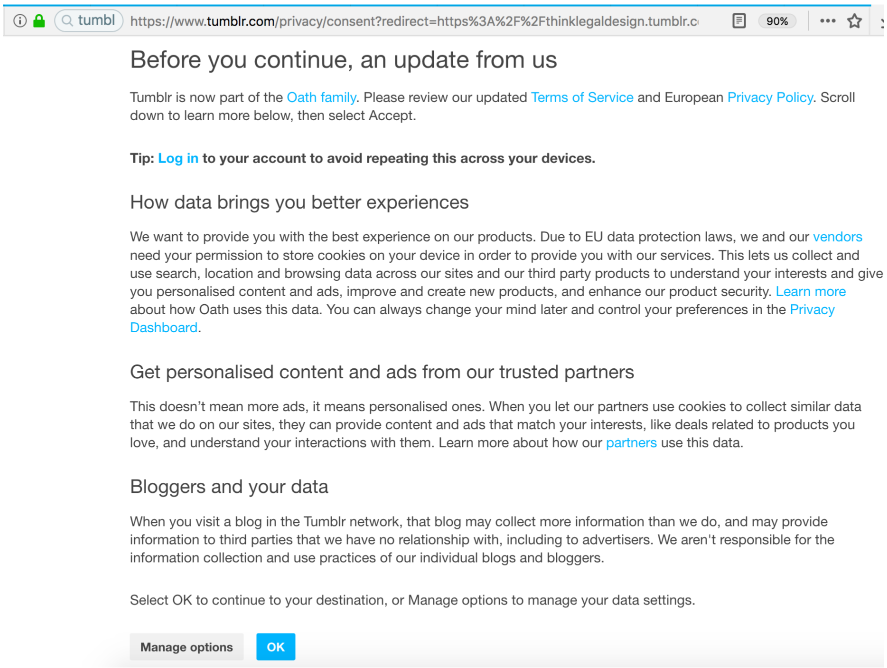
to the website, whereas she is still free to manage her options, but she must make an additional effort to do so.

2. If the user decides to manage her options, she lands on an additional screen (*see* Fig. 5.13) where it is explained (again) the necessity to store cookies on the device and the following positive consequences. At the end of the page, a bright blue "accept" button is conspicuously placed. If the user wants to manage the settings, she has to click on the "Manage" wording in the middle of the page: while buttons are made to be clicked, words can be clicked only if they are associated with an hyperlink, which is revealed only by hovering over it.

3. if the user decides, again, to manage the cookies, she lands on a page where every advertising partner is opted-in by default and the user must deselect them individually[11] (*see* Fig. 5.14). The total number of selections is 322 and there is no blanket button to disable every company at once. Needless to say, this activity takes minutes and even only the sight of the list of enabled cookies might discourage the user from engaging in such activity. In other words, the transaction cost (*see* Section 2.2) for the user might be too high for her to challenge the status quo, thus she might decide to take the more easily available choice: accept all the cookies and proceed with the navigation.

This is a real-world example that has been here analyzed from the point of view of the user. Although only one among many, this example strikingly shows the mechanisms that a company can exploit to lure users into specific, predetermined behaviors, such as sharing more personal data with advertising companies: users are given the possibility to change the settings and to choose the least invasive options, but these are made less prominent or are buried down behind buttons. Concretely, these privacy-preserving choices take more time and effort than accepting the privacy-corrosive options. In

[11] This is true for Tumblr. On other websites of the Oath parent company, like TechCrunch, there exists the possibility to deactivate them all at once



https://www.tumblr.com/privacy/consent?redirect=https%3A%2F%2Fthinklegaldesign.tumblr.c

Before you continue, an update from us

Tumblr is now part of the Oath family. Please review our updated Terms of Service and European Privacy Policy. Scroll down to learn more below, then select Accept.

Tip: Log in to your account to avoid repeating this across your devices.

How data brings you better experiences

We want to provide you with the best experience on our products. Due to EU data protection laws, we and our vendors need your permission to store cookies on your device in order to provide you with our services. This lets us collect and use search, location and browsing data across our sites and our third party products to understand your interests and give you personalised content and ads, improve and create new products, and enhance our product security. Learn more about how Oath uses this data. You can always change your mind later and control your preferences in the Privacy Dashboard.

Get personalised content and ads from our trusted partners

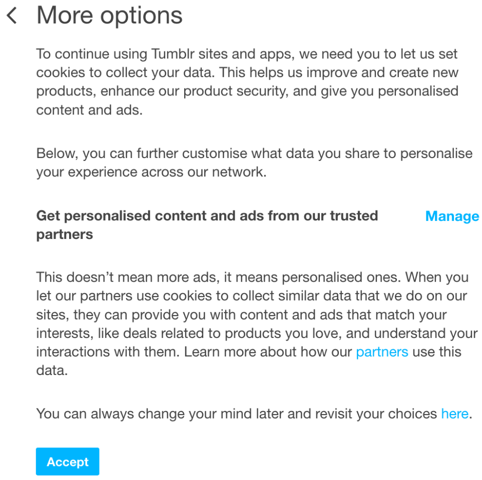
This doesn't mean more ads, it means personalised ones. When you let our partners use cookies to collect similar data that we do on our sites, they can provide content and ads that match your interests, like deals related to products you love, and understand your interactions with them. Learn more about how our partners use this data.

Bloggers and your data

When you visit a blog in the Tumblr network, that blog may collect more information than we do, and may provide information to third parties that we have no relationship with, including to advertisers. We aren't responsible for the information collection and use practices of our individual blogs and bloggers.

Select OK to continue to your destination, or Manage options to manage your data settings.

Figure 5.12: The cookie wall that Europeans found on every Oath's website (in this specific example, Tumblr) after the GDPR came into effect.



More options

To continue using Tumblr sites and apps, we need you to let us set cookies to collect your data. This helps us improve and create new products, enhance our product security, and give you personalised content and ads.

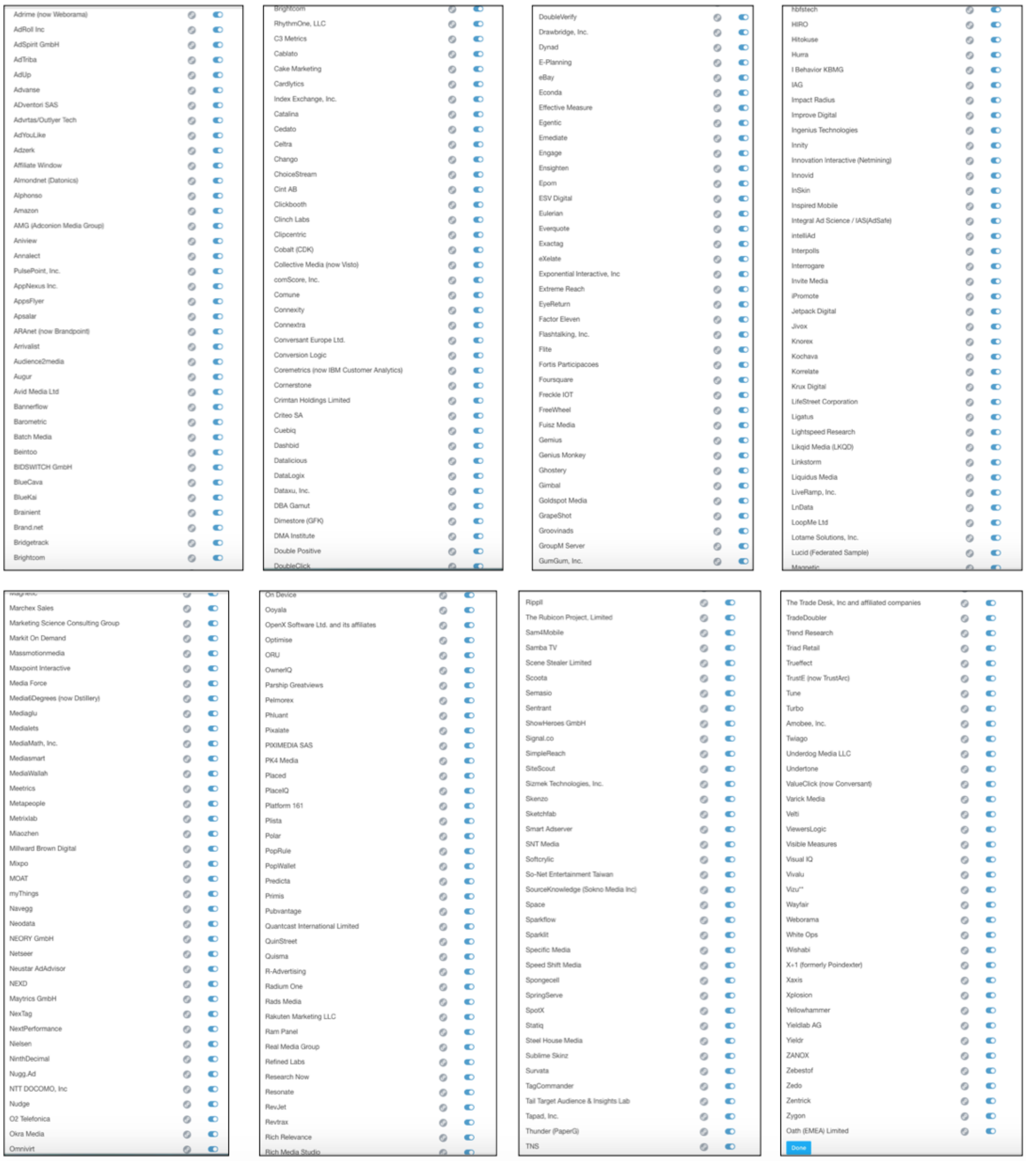
Below, you can further customise what data you share to personalise your experience across our network.

Get personalised content and ads from our trusted partners Manage

This doesn't mean more ads, it means personalised ones. When you let our partners use cookies to collect similar data that we do on our sites, they can provide you with content and ads that match your interests, like deals related to products you love, and understand your interactions with them. Learn more about how our partners use this data.

You can always change your mind later and revisit your choices here.

Figure 5.13: The page on which the user lands when she selects "manage options" on the previous page.



Adrime (now Weborama)
AdRoll Inc
AdSpirit GmbH
AdTriba
AdUp
Advanse
ADventori SAS
Advrtas/Outlyer Tech
AdYouLike
Adzerk
Affiliate Window
Almondnet (Datonics)
Alphonso
Amazon
AMG (Adconion Media Group)
Aniview
Annalect
PulsePoint, Inc.
AppNexus Inc.
AppsFlyer
Apsalar
ARAnet (now Brandpoint)
Arrivalist
Audience2media
Augur
Avid Media Ltd
Bannerflow
Barometric
Batch Media
Beintoo
BIDSWITCH GmbH
BlueCava
BlueKai
Brainient
Brand.net
Bridgetrack
Brightcom

Brightcom
RhythmOne, LLC
C3 Metrics
Cablato
Cake Marketing
Cardlytics
Index Exchange, Inc.
Catalina
Cedato
Celtra
Chango
ChoiceStream
Cint AB
Clickbooth
Clinch Labs
Clipcentric
Cobalt (CDK)
Collective Media (now Visto)
comScore, Inc.
Comune
Connexity
Connextra
Conversant Europe Ltd.
Conversion Logic
Coremetrics (now IBM Customer Analytics)
Cornerstone
Crimtan Holdings Limited
Criteo SA
Cuebiq
Dashbid
Datalicious
DataLogix
Dataxu, Inc.
DBA Gamut
Dimestore (GFK)
DMA Institute
Double Positive
DoubleClick

DoubleVerify
Drawbridge, Inc.
Dynad
E-Planning
eBay
Econda
Effective Measure
Egentic
Emediate
Engage
Ensighten
Epom
ESV Digital
Eulerian
Everquote
Exactag
eXelate
Exponential Interactive, Inc
Extreme Reach
EyeReturn
Factor Eleven
Flashtalking, Inc.
Flite
Fortis Participacoes
Foursquare
Freckle IOT
FreeWheel
Fuisz Media
Gemius
Genius Monkey
Ghostery
Gimbal
Goldspot Media
GrapeShot
Groovinads
GroupM Server
GumGum, Inc.

hbfstech
HIRO
Hitokuse
Hurra
I Behavior KBMG
IAG
Impact Radius
Improve Digital
Ingenius Technologies
Innity
Innovation Interactive (Netmining)
Innovid
InSkin
Inspired Mobile
Integral Ad Science / IAS(AdSafe)
intelliAd
Interpolls
Interrogare
Invite Media
iPromote
Jetpack Digital
Jivox
Knorex
Kochava
Korrelate
Krux Digital
LifeStreet Corporation
Ligatus
Lightspeed Research
Likqid Media (LKQD)
Linkstorm
Liquidus Media
LiveRamp, Inc.
LnData
LoopMe Ltd
Lotame Solutions, Inc.
Lucid (Federated Sample)
Magnetic

Magnetic
Marchex Sales
Marketing Science Consulting Group
Markit On Demand
Massmotionmedia
Maxpoint Interactive
Media Force
Media6Degrees (now Distillery)
Mediaglu
Medialets
MediaMath, Inc.
Mediasmart
MediaWallah
Meetrics
Metapeople
Metrixlab
Miaozhen
Millward Brown Digital
Mixpo
MOAT
myThings
Navegg
Neodata
NEORY GmbH
Netseer
Neustar AdAdvisor
NEXD
Maytrics GmbH
NexTag
NextPerformance
Nielsen
NinthDecimal
Nugg.Ad
NTT DOCOMO, Inc
Nudge
O2 Telefonica
Okra Media
Omnivirt

On Device
Ooyala
OpenX Software Ltd. and its affiliates
Optimise
ORU
OwnerIQ
Parship Greatviews
Pelmorex
Phluant
Pixalate
PIXIMEDIA SAS
PK4 Media
Placed
PlaceIQ
Platform 161
Plista
Polar
PopRule
PopWallet
Predicta
Primis
Pubvantage
Quantcast International Limited
QuinStreet
Quisma
R-Advertising
Radium One
Rads Media
Rakuten Marketing LLC
Ram Panel
Real Media Group
Refined Labs
Research Now
Resonate
RevJet
Revtrax
Rich Relevance
Rich Media Studio

Rippll
The Rubicon Project, Limited
Sam4Mobile
Samba TV
Scene Stealer Limited
Scoota
Semasio
Sentrant
ShowHeroes GmbH
Signal.co
SimpleReach
SiteScout
Sizmek Technologies, Inc.
Skenzo
Sketchfab
Smart Adserver
SNT Media
Softcrylic
So-Net Entertainment Taiwan
SourceKnowledge (Sokno Media Inc)
Space
Sparkflow
Sparklit
Specific Media
Speed Shift Media
Spongecell
SpringServe
SpotX
Statiq
Steel House Media
Sublime Skinz
Survata
TagCommander
Tail Target Audience & Insights Lab
Tapad, Inc.
Thunder (PaperG)
TNS

The Trade Desk, Inc and affiliated companies
TradeDoubler
Trend Research
Triad Retail
Trueffect
TrustE (now TrustArc)
Tune
Turbo
Amobee, Inc.
Twiago
Underdog Media LLC
Undertone
ValueClick (now Conversant)
Varick Media
Velti
ViewersLogic
Visible Measures
Visual IQ
Vivalu
Vizu**
Wayfair
Weborama
White Ops
Wishabi
X+1 (formerly Poindexter)
Xaxis
Xplosion
Yellowhammer
Yieldlab AG
Yieldr
ZANOX
Zebestof
Zedo
Zentrick
Zygon
Oath (EMEA) Limited
Done

Figure 5.14: An overview of the 322 active toggles that the user must individually deselect to opt out of targeted advertising on Tumblr

addition, it should be established whether the provision of default opt-ins is in line with the GDPR's provisions.

5.4 Conclusive Remarks

This Chapter has focused on the fundamental role that design can play in data protection to preserve and promote privacy in websites, apps, and devices. Thus, an analysis of the topic under multiple perspectives was offered: the newly introduced principles of data protection by design and by default of Article 25 GDPR were described in Section 5.1. These abstract principles have been coupled with concrete interventions to achieve transparency in privacy-related communication and to offer meaningful consent experiences. Design patterns for privacy are cornerstones of this approach: existing patterns were gathered, classified and described in Section 5.2. However, these solutions still constitute sparse and rare examples, therefore the hereby proposed classification can be subject to criticism and modifications: only the willingness of those actors in charge of the communication of data practices and the diffusion and experimentation of good practices will determine if such solutions will eventually become common patterns. One pattern that deserves scrupulous attention is the privacy icon design pattern, that will be thoroughly analyzed in the next chapter. Finally, privacy-eroding patterns were presented in Section 5.3, to demonstrate how the intention behind design choices is fundamental to accomplish goals that can be in line with legal and ethical principles, but also contrary to them.

Chapter 6

DaPIS: the Data Protection Icon Set

This chapter represents the focus of the present research: the creation and evaluation of a set of icons[1] for data protection, as a concrete realization of the privacy icon pattern presented in the last chapter. It starts with an introduction about the functions of icons as communicative devices, but also their potential limitations, especially when used in the legal sphere (*see* Section 6.1). Although icons are commonly regarded as elements that can convey meanings effortlessly across cultures, their effectiveness in fact depends on many (intrinsic and extrinsic) factors, for instance on the familiarity with the referents and the graphical representations. Such elements must be taken into account during the design and, especially, during the evaluation phase to properly reveal strengths and weaknesses of the icon set. The participatory design workshops organized to conceive, design and develop DaPIS, the Data Protection Icon Set, will be described in Section 6.2, whereas evaluation measures will be critically analyzed in Section 6.3.1. Then the several iterative phases of evaluation and consequent vetting of the icon set will be thoroughly illustrated, but also critically scrutinized, in Sections 6.4, 6.5, and 6.6. The

[1] Note that the term 'icon' assumes in this Chapter a different meaning than the peircean one, introduced in Section 4.4. Here the term refer to its HCI definition, as a pictographic representation.

chapter ends with a proposition for an icon set that is based on PrOnto, the privacy ontology presented in Chapter 3, and that can act as navigation aid in lengthy privacy policies. Limitations on the effectiveness of DaPIS and recommendations for its widespread adoption conclude this chapter.

With the coming into effect of the GDPR, the theoretical discussion and the provision of practical examples, as well as evidence on how to produce, evaluate, and use icons for data protection, have become timely and needed (*see* also [30]). Icons that accompany the different thematic sections of privacy policies gain even more relevance because the quantity of information about data practices that must be disclosed increases per Articles 13-14. As a consequence, privacy notices' length also increases: information architecture and visual elements can greatly contribute to ease the navigation of the documents. As it will be argued later, not only data subjects, but especially those actors that consult these statements more often such as lawyers, consumers associations and auditors, will benefit from icons as information-markers.

6.1 Icons for the Law

In the last decade, "privacy icon sets" have started to appear as private or company-lead initiatives, but they have not met widespread adoption. Since the market seems to have failed to take the initiative upon itself [91], the GDPR suggests the use of visualizations (Recital 58) and specifically of icons (Article 12.7) to enforce the principle of transparency, namely to provide "in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing". This approach is aimed at reducing excessive amounts of written information and research in this direction is encouraged by the Article 29 Working Party [30] for an evidence-based approach that can inform the application of icons in this context. Similarly, the Consumer Rights Directive [227] has included an Optional model that provides a standard format with icons to summarize and illustrate contract terms. Both the Regulation and the Directive do not set icons as obligations, but rather suggest their possible adoption to data controllers, in the first case, and sellers, in the second case. Although eventually it will be the role of the European Commission to adopt delegated acts to give directions on the creation of these icons, the need for expert advice is emphasized in Recital 166 GDPR and in the Guidelines on Transparency [30]: this chapter intends to be a contribution to the preparatory work.

6.1.1 On the Nature of Icons

Icons are attractive communicative devices because they can be easily recognized, processed and memorized. They can serve as memory devices and help in the classification of content [194]. Icons are deemed to communicate in a nonverbal manner quickly, concisely, and across linguistic and cultural differences [165]. However, one must be cautious about these claims (*see* e.g. [84]). For instance, familiarity with the icon is a highly relevant dimension to preview ease of access to memory and time of recognition, and has a twofold dimension: it is affected by previous experience with the graphical symbol, but also with the symbol's underlying concept.

Icons can be also highly diverse in terms of their semantic distance, which determines their interpretability and ease of learning [165]. They can be placed on a continuum that ranges from resemblance icons (that depict objects), to exemplar icons (that portray an individual of a class), to symbolic icons (that convey a concept on a higher level of abstraction) to arbitrary icons (that have no relationship to objects or concepts) [194]. Moving towards the end of the spectrum, the semantic transparency of the symbols shrinks [206] and their meaning must be learned rather than deduced [150]. Moreover, the function assigned to a graphical symbol by its designer can be different from the meaning attributed to it in practice, i.e. there can be misalignments between the designers' intentions and the sense-making activity of the user (see Section 4.5).

For new exposures, ease of identification also depends on the icon's concreteness, which is the extent to which real objects, materials, or people are depicted [165]. This effect diminishes as users gain experience over the icons, though. An additional aspect that must be considered is complexity, i.e. the amount of icon's details, that influences search activity, but not identification. Finally, when creating a set of icons, attention should be devoted to the degree of discriminability of one icon from the others of the set and to coherence across set elements [84]: the same graphical symbol should be used consistently to signify the same meaning, but it also should be enough distinctive from the others to be easily identified.

By considering these dimensions illustrated in Fig. 6.1 and the visualizability of the underlying concept [299], an icon's cognitive effectiveness (speed, ease, and accuracy of interpretation) [206] can be roughly estimated even before an empirical evaluation. If these dimensions are not carefully considered, there is risk that users will process the visual representations more slowly, with more difficulty and with less success compared to written text. By doing so, obscurity in lieu of transparency would be achieved and the very goal of the icons would not be attained.

6.1.2 On the Role of Legal Icons

Probably the most common and easily accessible example of universally understandable iconic language in the legal domain (i.e. "graphic law" [213, p. 780]) is represented by the code of the road. The impact of traffic signs on human behavior mainly occurs, indeed, through visuals that completely substitute verbal utterances. This is why the symbols of the highway code should be unambiguous: easily decipherable and immediately understandable by all citizens [289]. However, such symbols are not universally recognized and correctly interpreted because of their semantic transparence (*see* the example in Fig. 6.2). The almost total absence of text is rather deemed a proof of sedimentation of a specific area of legal knowledge. This is possible because the road signs visually codify meanings that have been learned and internalized by the whole community: drivers are considered and treated as experts in the law even though they are not. Such sedimentation was

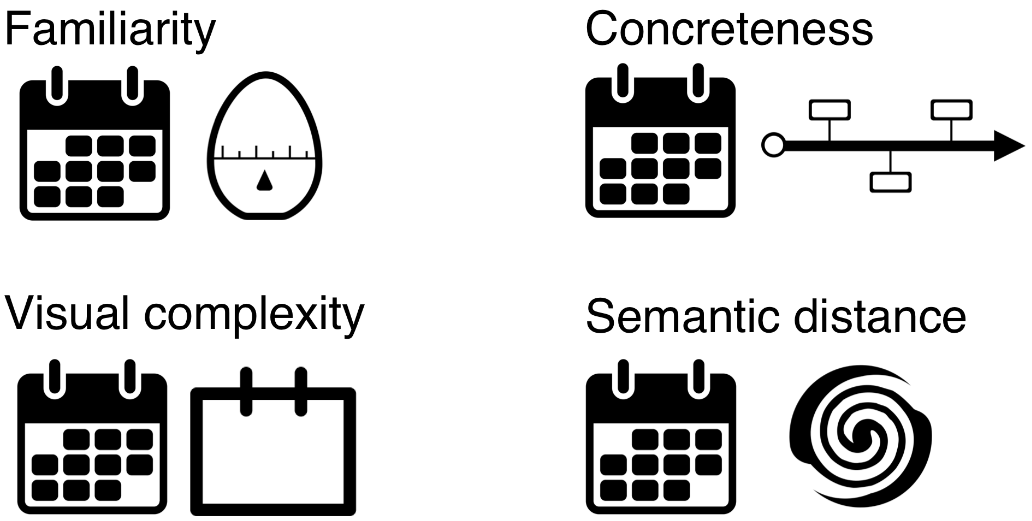


Figure 6.1: Examples of icons for the concept of "time" according to the 4 relevant dimensions for icons classification and recognition: familiarity, concreteness, visual complexity, semantic distance. On the left-hand side is shown a prototypical example of the category, whilst on the right-hand side is shown an example on the other hand of the spectrum

achieved through constant use, converted into traffic regulations, and through international harmonization and standardization of the norm, e.g. following international conventions, e.g. [282].

Indeed, icons have limited self-explanatory nature [155]: decoding these pictograms requires context and learned knowledge (e.g. cultural knowledge). Icons that convey abstract meanings, such as data practices, might not be universally understood if they are not accompanied by some textual explanations [295]. Usability tests [155, 239, 159] show that "critical confusions" [298], namely misinterpretations opposite to the intended meaning, are possible due to multiple reasons: misalignment between designers' intentions and users' expectations on the icon meaning and differences in individuals' level of education, age, and cultural background. This matter assumes great relevance if individuals take legally-binding decisions based on the visualizations, such as entering into a contract with a service provider or giving consent to certain data practices. Indeed, legal meaning encoded in pictures

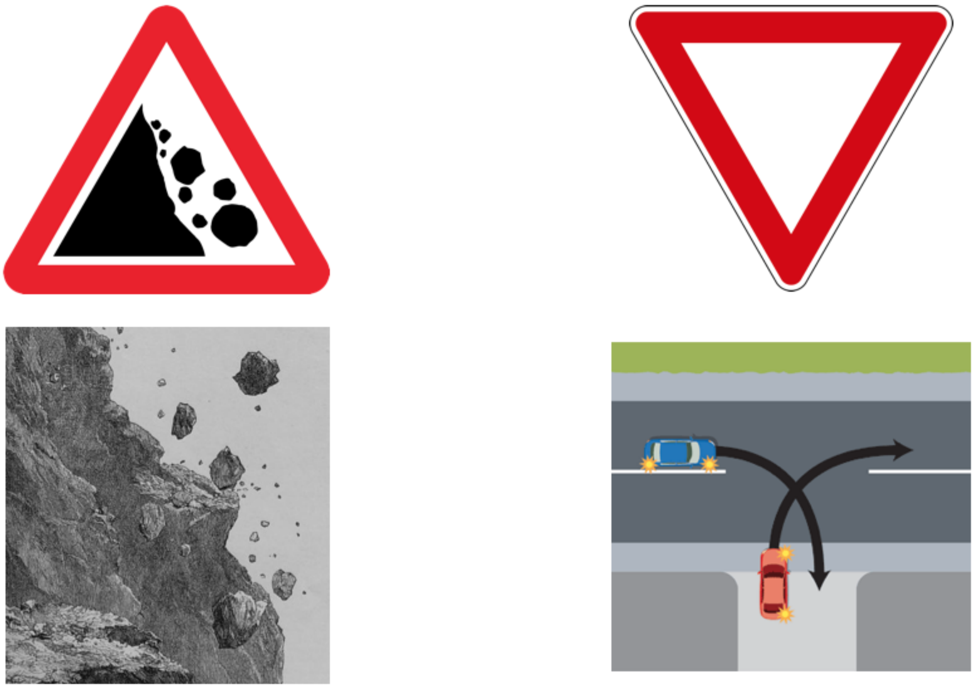


Figure 6.2: Two traffic signs of different indexical nature. On the left a symbol that transparently represents the concept of falling rocks. On the right, the arbitrary symbol for the concept of give way, showing no pictorial adherence to the underlying concept.

is open to multiple interpretations [46] and there are serious concerns that pictograms do not represent legal concepts and norms in terms of details and adequateness as words would do. Thus, businesses do not reasonably want to risk misunderstanding of the pictorial representations or oversimplification of their privacy terms, because this might cause liability issues [144].

As recalled earlier, familiarity is a critical component to determine ease of recognition of an icon. This is why good practice for icon design is to rely on an established visual vocabulary [150]. However, this proves difficult in the legal sphere because there exist only a few, overly preponderant, law-related symbols, such as the scale and the gavel. As for what concerns data protection, only a few symbols around (cyber)security are well-known, such as the shield. Different is the case of technology-related visuals, since the widespread use of graphical user interfaces has favored the creation of mental references between a number of icons and their functions (e.g. a pencil for the edit function). These conventions have been reused in the design of DaPIS.

In addition, familiarity with the concept underlying the icon plays a fundamental role: if the concept is unknown to the interpreter, as it is generally the case with legal matters, then the icon must possess a low level of arbitrariness to easily shed light on its underlying meaning. An additional difficulty is posed by the fact that legal concepts are usually abstract in nature, so it becomes even more difficult to visualize and consequently depict them.

Such characteristics also challenge classical evaluation methods, which are mainly suited to determine the comprehensibility of graphical symbols whose referent is known to the user (*see* also Section 6.3). Differences of comprehension rates are to be expected, because they depend on the intrinsic icon's characteristics, i.e. familiarity, concreteness, and semantic distance, but also on the characteristics of the person that interprets them, i.e. culture, age, etc. Finally, researchers underline the importance of the provision of contextual cues that mirror the actual usage situation of the icons (ecological validity [165]) to support the sense-making process of individuals. Without such precautions, low recognition scores might falsely indicate that more design and test work is necessary [300].

6.1.3 On the Role of Data Protection Icons

Icons belong to the category of the visual representation patterns, which help to explain, complement, and disambiguate the (legal) text. "Companion icons" are "graphic symbols that represent the meaning or function of the textual element they accompany" [136, p. 26]. They help readers to search and find relevant information quickly, especially in long and undifferentiated texts such as privacy policies: they act as a visual index that helps the data subject to find the topic she is looking for. Thus, icons can in principle highlight and quickly communicate the key aspects of the privacy practices of an organization [134]: it is commonly believed that "a privacy icon[2] is worth a thousand-word policy[3]".

However, it is a common misconception in the legal sphere that icons, or visual elements more in general, should substitute words and text completely [91]. Rather than substituting the legal text, data protection icons can integrate it and act as information markers, namely to help the reader to quickly navigate or skim through long texts [230]. Used in combination with a structured layout, they can help data subjects to quickly find specific information items and, thus, to exercise strategic reading. They can also attract the attention of the reader, fight information fatigue, and help to memorize information. In principle, they can even provide a short summary of the privacy practices at a glance. It is however questionable whether they should also provide a judgment on the fairness of terms (*see* also Section 6.1.4 and [239, 125]).

In the interpretation of the Article 29 Working Party, the icons are meant

[2] The literature generally refers to icons depicting concepts related to data practices as "privacy icons". However, they mostly represent concepts of data protection, thus the term is inexact. In the present report, the expression "data protection icons" will be preferred.

[3] *See* the "privacy icon pattern" on `https://privacypatterns.org` and `https://privacypatterns.eu`.

to enhance transparency by reducing the extreme amount of information and, upon standardization, to be used across the continent as universal shorthand for that information [30]. However, reducing the complexity and the potentially infinite combinations of linguistic terms into a limited set of icons is impossible (for critical remarks, *see* also [203]). Icons are probably the visual means that can be supposedly more easily accepted by the legal world than other categories of visual elements, since iconic pictograms have already met widespread adoption to depict traffic laws and have been successfully used to quickly display Creative Commons' licenses to grant copyright permissions (*see* also [142]).

6.1.4 Previous Work on Data Protection Icons

There have already been some attempts to design a "visual language for privacy data rights" [243]. A few privacy-related icon sets already exist [257, 199, 122, 159, 208, 125, 102, 281] and vary deeply in nature:

1. as for what concerns the types of information that they represent: this dimension depends on the typology of service (e.g. websites, social networks, e-mails), on the legal framework, on whether the fairness of data practices is displayed, as well as on researchers' choices. Great variety exists in terms of quantity of elements in the set and level of detail, but there are also some recurring categories, e.g. type of data, processing and collection purposes, time of storage, sale of data.

2. whether or not they represent a legal assessment about the fairness (in terms of adherence to users' expectations or of legal compliance) of the terms they represent: for instance, Mozilla [208] and TRUSTe with Disconnect [281, 241] designed icons that display straightforwardly whether a website collects and processes data in a manner that is foreseeable by the data subject.

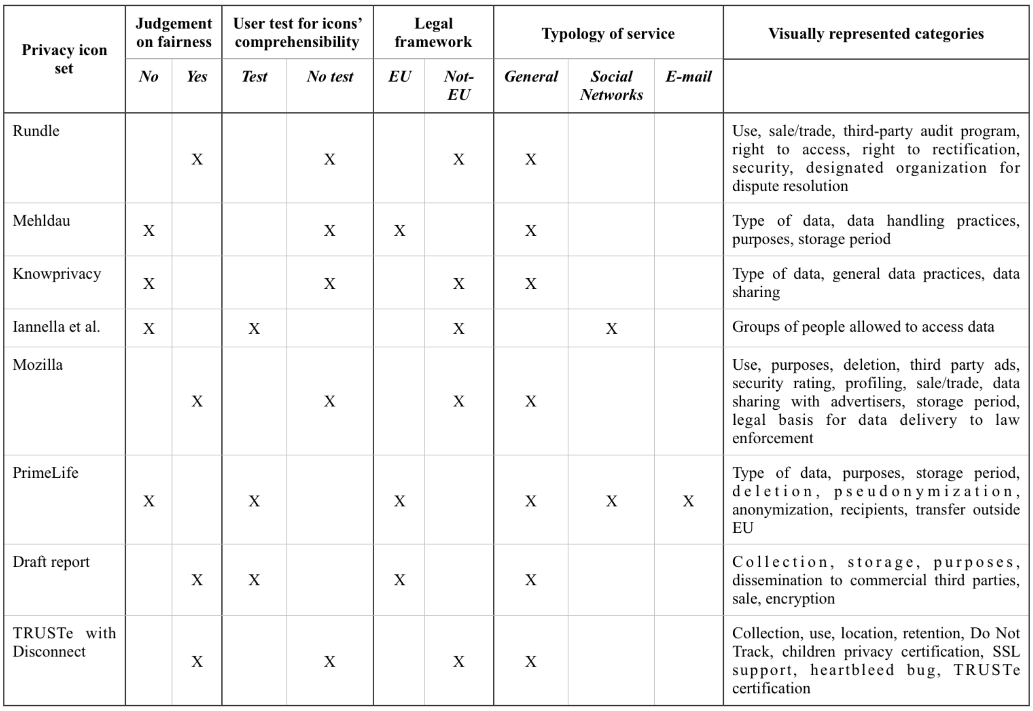
3. whether or not they have undergone a user study about their ease of comprehension or other dimensions (e.g. legibility): as elaborated

previously, well thought-through design translates into the adherence between designers' intentions and users' mental models. In the case of privacy icons, this aspect bears considerable relevance because organizations might not want to risk misunderstandings, that could cause liability issues [144]. User research has, thus, the function of confirming or rejecting the hypotheses formulated during the design stage: for instance, usability testing in the PrimeLife project [155] revealed that users' characteristics like age, education, as well as cultural and educational background play a role in the interpretation of privacy-related pictograms and can thus undermine the supposedly universality of such symbols. However, the literature review that we carried out in [254] highlighted that only a small number of studies [159, 155, 239] on icons' comprehensibility were carried out.

4. the regulatory framework they refer to (EU or USA): most icon sets were designed for a US audience and with reference to those data practices that appear more relevant in an American context. Noteworthy, there are two approaches within a European perspective that need to be mentioned for their relevance to the present research: the Primelife project [125] and the icons in the Draft Report on the Regulation Proposal [102]. These two attempts will be critically examined in the following sections, also because they provide the opportunity to discuss the evaluation methods applied to data protection icons.

Fig. 6.3 summarizes the belonging of each icon set to the selected criteria.

In general, attempts to consider the context where the icons would appear have been scarce, with the exception of TRUSTe with Disconnect [281, 241] and the Draft report [102]. In the latter case, however, it is the precriptive character of the specific context suggested (i.e. a table) that constitutes one of the main critical points, as will explained below in Section 6.1.4.2.



Privacy icon set	Judgement on fairness		User test for icons' comprehensibility		Legal framework		Typology of service			Visually represented categories
	No	*Yes*	*Test*	*No test*	*EU*	*Not-EU*	*General*	*Social Networks*	*E-mail*	
Rundle		X		X		X	X			Use, sale/trade, third-party audit program, right to access, right to rectification, security, designated organization for dispute resolution
Mehldau	X			X	X		X			Type of data, data handling practices, purposes, storage period
Knowprivacy	X			X		X	X			Type of data, general data practices, data sharing
Iannella et al.	X		X			X		X		Groups of people allowed to access data
Mozilla		X		X		X	X			Use, purposes, deletion, third party ads, security rating, profiling, sale/trade, data sharing with advertisers, storage period, legal basis for data delivery to law enforcement
PrimeLife	X		X		X		X	X	X	Type of data, purposes, storage period, deletion, pseudonymization, anonymization, recipients, transfer outside EU
Draft report		X	X		X		X			Collection, storage, purposes, dissemination to commercial third parties, sale, encryption
TRUSTe with Disconnect		X		X		X	X			Collection, use, location, retention, Do Not Track, children privacy certification, SSL support, heartbleed bug, TRUSTe certification

Figure 6.3: Classification of existing privacy icon sets, ordered according to the year of their first publication. Legend: Rundle [257], Mehldau [199], Knowprivacy [122], Iannella et al. [159], Mozilla [208], Primelife [125, 155, 106], Draft report [102], TRUSTe with Disconnect [281, 241]. Adapted from [254].

6.1.4.1 The PrimeLife project

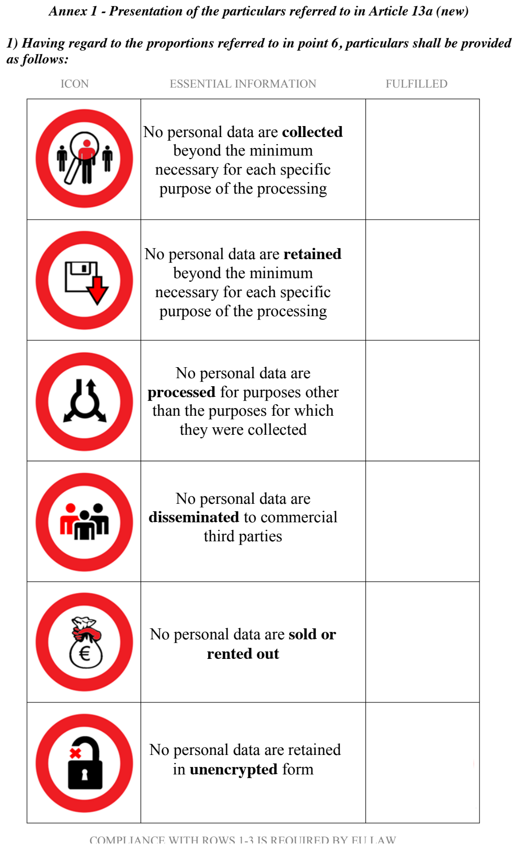
The PrimeLife project [125] is notably the most structured attempt to design and evaluate icons for data protection in the European context. The first icon set produced during the project comprises symbols representing data processing steps of various kinds, types of data, processing purposes and categories of recipients (the latter only for social network). However, most of the icons produced during this first step of the project were discarded during the testing phase. Although the user study [107] highlighted how visual vocabulary depends on culture, therefore calling for intercultural user audiences, it is difficult to determine its ecological validity since it did not provide much context to support the sense-making of the test participants. In general, icons with labels were better understood than the same icons without labels. No specific numbers about the results are provided, but it comes at no surprise that the icons that scored best (i.e. medical data, payment data, storage, deletion, etc.) refer to more concrete and more familiar referents, whilst the less recognized icons (such as anonymization, user tracking, etc.) depict less familiar and concrete concepts.

Then, another test with a wider audience was conducted [125]. It was asked either to decide between a few possible alternatives or to rate icons according to their comprehensibility, clearness, and feasibility. Participants could even add comments on their own and elaborate on reasons for critique or approval of the icons [155]. Some principles that emerged were: simplification of the elements is crucial, as well as uniformity of the design styles. The PrimeLife's researchers end with a negative note: given the low results in both user studies, only a few icons were deemed appropriate to be included in the final icon set: third party sharing, storage period, third party tracking and behavioral targeted advertising (plus three icons about data disclosure in social network sites). At a careful analysis, it seems that such icons depict concrete or familiar concepts. As highlighted earlier, however, results on icons' comprehensibility mainly depend on their semantic distance and familiarity, which are dimensions to be taken into consideration during evaluation. Many other processing steps, data types and recipient groups were deemed too hard to illustrate or recognize. Although the goal of the PrimeLife project was the creation of icons for an interface, it seems that no test in context was carried out, which could have sparked higher results.

6.1.4.2 The Draft Report on the Regulation Proposal

During the parliamentary discussion about the GDPR, a table with 6 icons (*see* Fig. 6.4) was proposed to summarize the main data practices of a data controller. Such table appeared in an Annex to the 2013's Draft report of the LIBE Committee on the Regulation proposal [102]. The display of such icons would have constituted a legal obligation for websites, were the amendments approved. Trace of this proposition can be found in the GDPR's call for icons. Instead of "neutrally" translating privacy notions into visuals, these icons symbolize assessments about the website's compliance with six basic data protection guarantees. The assessment about compliance is derived by the combination of the symbols on the left-hand column and the evaluation indicators (i.e. a cross in a red circle or a tick in a green circle) on the right-hand column. The icons signify fair data practices: e.g. no data collection nor data retention beyond the minimum necessary, no data dissemination to commercial third parties, etc.

The comprehensibility test carried out on these icons [239] shows some shortcomings as for what concerns its ecological validity, as the research's author himself acknowledges. The first part of the test asked for the icons' meaning without providing any contextual reference. However, the icons were explicitly designed to appear next to the corresponding textual explanations, thus assessing their meaning in isolation does not constitute an helpful judgment. The second part of the test consisted in a matching task between icon and correspondent textual explanation, where multiple associations, as opposed to one-to-one associations, were allowed in order to determine if multiple matches, meaning confusion, would occur. The test also produced important results: for example, the use of the combination of one icon on the



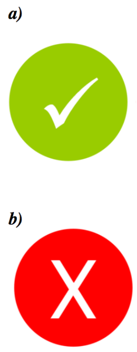
Annex 1 - Presentation of the particulars referred to in Article 13a (new)

1) Having regard to the proportions referred to in point 6, particulars shall be provided as follows:

ICON	ESSENTIAL INFORMATION	FULFILLED
	No personal data are **collected** beyond the minimum necessary for each specific purpose of the processing	
	No personal data are **retained** beyond the minimum necessary for each specific purpose of the processing	
	No personal data are **processed** for purposes other than the purposes for which they were collected	
	No personal data are **disseminated** to commercial third parties	
	No personal data are **sold or rented out**	
	No personal data are retained in **unencrypted** form	

COMPLIANCE WITH ROWS 1-3 IS REQUIRED BY EU LAW

(a) Privacy icons and their description in a table



(b) Graphical symbols to signal the fulfillment of the conditions laid down in the second column

Figure 6.4: The tabular format proposed in Annex 1 of the Draft report on the Proposal for the GDPR [102] for standardised information policies. The first column contains privacy icons, the second column contains the conditions represented by the icons, while the third column must be filled by the data controller with either one of the graphical symbols of Fig. 6.4b, depending on whether the condition is fulfilled.

left column and the cross for a double negation was deemed hard to understand, because it would mean, e.g. “It is not the case that it is not the case that personal data are disseminated to commercial third parties”. Besides, if the icons signify a negative statement, this would have been more easily understood through the use of a diagonal bar, because it would resemble our shared visual mental model for negations. The author of this study recommends to not use icons to make statements, but rather to indicate areas: e.g. signal to users in which part of the privacy notice they can find the description of the practices about data dissemination to commercial third parties.

Finally, as briefly mentioned earlier, even the table where the icons are displayed can be considered more as a nuisance than as a helpful suggestion for concrete applications. Although the tabular format represents one of the rare, but necessary, existing contextual indications for the icons, it also greatly limits their use. For example, it is inconceivable for such a format to be employed on small screens.

6.1.4.3 Neutral Representation or Assessment on Fairness?

As pointed out by the two last examples [125], we also share the idea that icons should have a headline function, rather than make a statement about the fairness of processing, in order to reach global acceptance. It can be argued that it would be more meaningful for data subjects to be provided with a visual summary of the risky or less lawful practices conducted by an organization on their data (i.e. a rating), in order to support their decision-making, e.g. if to use a certain service or head elsewhere. However, such an approach encounters the problem that a decision about the lawfulness and fairness of certain practices should be taken and it is questionable who should take it and on the basis of which principles. Moreover, such an approach would be probably opposed by many organizations, since the GDPR does not impose an obligation upon controllers: icons rating lawfulness would therefore be very difficultly to be widely adopted. These are the reasons why the present research adopts a more neutral approach, namely depict notions of data protection to act as information markers, whereas it will be the individual to decide for herself whether to engage with a certain service. Finally, depicting concepts is also the most suitable integration with an ontological formalization of legal knowledge.

6.1.4.4 Emerging GDPR Icons

A few examples of the use of icons in privacy communication have also recently emerged and will be here compared to DaPIS. In the French area, there have been two notable attempts to design icons for data protection after the GDPR's Article 12. The Association Privacy Tech has recently published a set of icons [276] depicting types of personal data, data storage duration, recipients, a few purposes, and extra-EU data sharing. However, there is no mention about the design process and possible evaluations. Another recent initiative is represented by [3], that has released 73 icons describing types of data, processing purposes, and other categories at a very fine-grained level. Again, no reference to the methodology for the design and evaluation is provided.

Special mention deserves Juro's privacy policy [11] (*see* Fig. 5.3) designed by Stefania Passera, a legal design pioneer who has been widely cited in Chapter 4 and who has made extensive research on successul design patterns applied to legal documents [136]. Information architecture and visual cues, among which icons, have been experimentally deployed to make Juro's privacy information compliant with the GDPR's transparency obligation. Expert advise from privacy professionals was sought during the design phase, but no user study on icons' comprehensibility was carried out. Further research could compare the efficacy of DaPIS with Juro's icons to determine whether participatory design methods are able to align designer's intentions and users' interpretations with more success if compared to one only designer's choices.

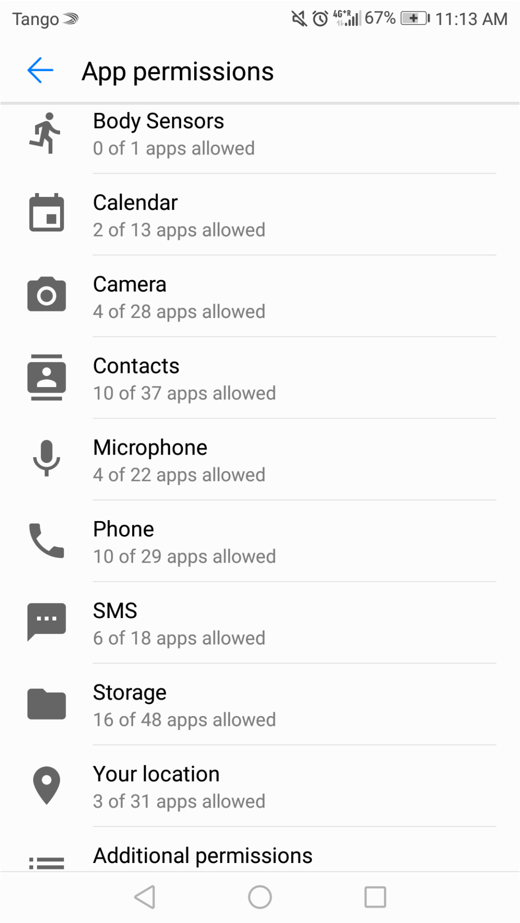
6.1.4.5 Icons for App Permissions

To complete the picture, it is necessary to introduce and critically discuss the privacy-related icons that have actually met worldwide adoption. In this context, app permissions provide the best examples of the use of icons to signal to users the types of data collected (*see* fig. 6.5). Such an approach is officially endorsed by the Article 29 WP [23] with the aim to raise awareness of users about data processing and is deemed appropriate to convey information on small screens in combination with a minimal amount of information e.g. through layered notices[4].

There are four main moments [260] in which access to data is made relevant and visible to users on mobile devices through graphical symbols. The first case is represented by the permissions asked by the app at the moment of download (i.e. at setup notices). The same permissions are manageable in the settings at any moment that the users can find when they are actively looking for them (i.e. on demand notices), divided per app or per type of data requested (*see* Fig. 6.5). Additionally, contextual just in time notifications are shown at the moment of collection of data, such as when an app requests access to photos and contacts (*see* Fig. 6.6). Finally, a persistent notification can be shown to raise user's awareness while a data practice is ongoing, e.g. the geolocation icon on the smartphone status bar.

The supervisory authorities recommend icons to be "meaningful, i.e. clear, self-explaining and unambiguous" [23, p. 24] and also suggest consumer testing to make sure that any form of communication is "understandable to users without a technical or legal background" [23, p. 24]. However, the icons currently employed in mobile devices exclusively depict types of data which are easier to visualize (e.g. the image icon) or have already secured a place in the visual vocabulary of any common user (e.g. the geolocation icon). As explained in the following, it is much more difficult to create self-explaining icons for other types of notions pertaining to data protection.

[4] Identical recommendations are provided to manufacturers of connected vehicles to easily and quickly signal to the drivers the data practices active on the vehicle [185]


Tango

67% 11:13 AM

App permissions

Body Sensors
0 of 1 apps allowed

Calendar
2 of 13 apps allowed

Camera
4 of 28 apps allowed

Contacts
10 of 37 apps allowed

Microphone
4 of 22 apps allowed

Phone
10 of 29 apps allowed

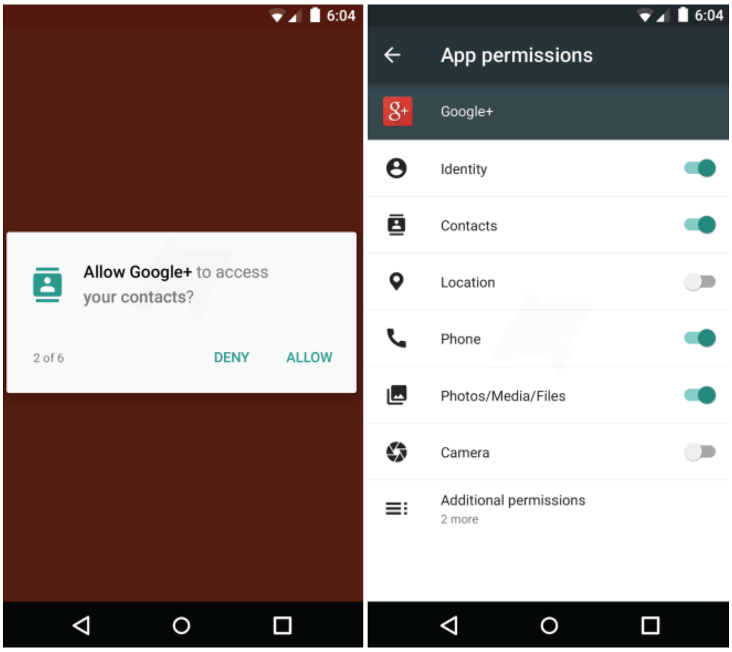
SMS
6 of 18 apps allowed

Storage
16 of 48 apps allowed

Your location
3 of 31 apps allowed

Additional permissions

Figure 6.5: Icons in the app permissions on Android. On the left, just-in time notice at the moment of data collection and on the right, on-demand notice in the app's settings



Allow Google+ to access your contacts?

2 of 6

App permissions

Google+

Identity ●
Contacts ●
Location ○
Phone ●
Photos/Media/Files ●
Camera ○
Additional permissions
2 more

Figure 6.6: Icons in the app permissions on Android

6.2 The Design of DaPIS

The analysis reported in the previous section highlighted some limitations and issues that the present research aims to tackle, by proposing a methodology to design a privacy icon set [223] that represents core concepts of European Data Protection Law. Notwithstanding the dismal results of the usability tests on previous privacy-related icons, that led the researchers to discard the majority of icons, it is hereby argued that previous experiences can be improved and icons to convey legal meanings can be more successful, based on results in comparable research [230]. Therefore, the present chapter presents DaPIS: the Data Protection Icon Set. The project described in the following employs experimental, human-centered design practices and semantic web technologies to satisfy GDPR's legal requirements about transparent information provision.

The approach proposed in the following has the final aim to semi-automatically display the icons in correspondence of the matching privacy terms (*see* Appendix A). The underlying hypothesis, derived from [230, 136], predicts that this can make statements on specific topics in privacy policies easier to find and understand. Clickable icons can also be employed to signify a data subject's explicit consent to certain practices [31]. A methodology based on three steps is here proposed (*see* also [223]):

1. Formalization of legal knowledge: Sect. 6.2.1 defines the objects of representation;

2. Participatory design methods: Sect. 6.2.2 describes the design process of DaPIS over some participatory design workshops;

3. Evaluation: Sect. 6.3 describes methods and measures for the assessment of DaPIS.

,

6.2.1 Formalization of Legal Knowledge

None of the other data protection icon sets are not based on a systematic formalization of knowledge, but rather focus on data types and a handful of processing operations (*see* Section 6.1.4). Moreover, the GDPR introduces several new information items that must be presented to data subjects to enforce the principle of transparency (*see* Artt. 13-14), such as some rights, for which no graphical representation has been proposed yet.

The GDPR provides a European legal framework that defines concepts of data protection, relations among them, and a common vocabulary to describe them. Since the Regulation demands "machine-readable" icons[5] if presented on electronical means (Art. 12.7), DaPIS is modelled on the computational ontology PrOnto (*see* Section 3.4). When a privacy policy is marked up with tags linked to the ontological instances, its semantic content can be described in a machine-readable manner. Icons can be associated to the concept

[5] Although the GDPR does not provide a definition of machine-readable, Recital 21 of Directive 2013/37/EU17 defines it as "a file format structured so that software applications can easily identify, recognize and extract specific data, including individual statements of fact, and their internal structure". *See* also [30]

they represent and, hence, be semi-automatically summoned by the semantic tags[6]. The description of legal information in a machine-interpretable format also allows automated reasoning on the legal texts, e.g. to draw inferences to match expressions in natural language to the corresponding ontological instance. Finally, an ontology is independent from language, which counts as an additional strength: the same icon can be provided for text spans expressed in different languages, that however refer to the same ontological concepts, whilst correspondent labels in different languages can be provided for the same icon.

6.2.2 Participatory Design Workshops

The risk of misalignments between designers' intentions and the sensemaking activity of individuals oriented the research towards participatory design methods (*see* Section 4.2) for the creation of the icon set. Indeed, previous research has found that it is arduous for experts to think like non-experts and, thus, symbols created by the target audience are more likely to be correctly interpreted by other members of the target audience, since they share similar mental models and cognitive profiles [57].

Collaborations among experts in different areas and laypeople can, on the one hand, leverage on the multiple skills and knowledge of the different stakeholders involved. On the other hand, this reduces the chances of personal bias because reciprocal understanding is deliberately sought [42]. Diverse mental models and visual vocabularies derived by different backgrounds and experiences are thus considered. In the four participatory, multi-stakeholders' workshops organized to create DaPIS, multi-disciplinarity was a critical element, thus motley working groups were formed. Each background represents an asset [255]: (1) legal experts explain data protection concepts and exemplify their meaning; (2) computer scientists or participants with similar

[6] Provided the development of such a tool, which was not the goal of the present project, though.

Superclass	**Class**
Personal data types	Original personal data Derived personal data Inferred personal data
Agents' roles	Data subject Data controller Data processor Supervisory authority Third party
Processing operations	Copying Pseudonymization Anonymization Direct marketing Automated decision-making Profiling Encryption Transfer of personal data to third countries
Data subject's rights	Right to be informed Right of access Right to rectification Right to erasure Right to withdraw consent Right to data portability Right to restriction of processing Right to object to processing Right to lodge a complaint
Processing purposes	Research purpose Statistical purpose Purpose of information security Purpose of provision of the service Purpose of service enhancement Marketing purpose Profiling purpose
Legal bases for processing	Consent Legal obligation Vital interest Public interest Legitimate interest Contract

Table 6.1: Conceptual cores of the GDPR ontology, on which DaPIS is based

backgrounds have the technical expertise to understand and explain technical notions included in the data protection law; (3) graphic designers and other professionals from visual disciplines know the techniques and tools to produce functioning visualizations for the intended audience and the intended medium; and lastly (4) laypeople add non-expert, but also non trivial views and knowledge to the design process, for instance about the visual conventions they are familiar with.

In participatory design cycles, multiple ways of representing the same concept are collaboratively examined. In our experience, the level of detail of the visualization was source of discussion and represented the main tension between precision of representation (favored by lawyers to avoid oversimplification) and simplicity (endorsed by designers to attain usability). Unlike other disciplines, design thinking does not aim for a prescriptive theory to generate a single "right" image or layout, because it could be unsuitable for individual needs. Preferable is a collaborative and creative process that tends towards a visualization that "works", given information type and goal of the design [42]. This is why, discussions about low fidelity prototypes in small and then bigger groups is encouraged, as well as intercept interviews with individuals that have not participated in the creation phase. By doing so, shortcomings and strengths of the proposed ideas are identified before the actual creation of high fidelity icons and chances of failure at later stages (e.g. during the evaluation phase) can be minimized.

The DaPIS was developed during four participatory legal design workshops that are briefly described in the following Sections.

6.2.2.1 The Design of a First Icons' Subset

A first, exploratory workshop [183] was held at the Legal Design Lab [140] of Stanford Law School in July 2017. The workshop was structured around the design cycle, over 6 hours. After a presentation of the GDPR and its call for icons, previous attempts to create data protection icons were reviewed to select promising elements composing a shared visual vocabulary.

Interdisciplinary groups were formed to create icons for the following classes of concepts (*see* Table 6.1): (a) data types (e.g. processed data, inferred data, etc.), (b) agents' roles (e.g. data subject, data controller, etc.), (c) processing operations (e.g. copying, transfer of data outside of the EU, etc.) and, finally, (d) the right of access and the right to data portability. At this stage, workshop participants worked with pencils and papers to explore different possible visualizations of such concepts and to generate low-fidelity visuals. These prototypes were tested internally, with the entire workshop group for a critical review, and externally, by conducting five to ten intercept interviews on Stanford campus to get early, unstructured feedback. Eventually, a workshop plenary discussion identified those iconographical elements that deserved to be kept, those that deserved further elaborations and those that needed to be abandoned. A graphic artist then collaborated with the researchers to render the draft icons digitally and to harmonize their style, also thanks to additional, recursive small sample testing conducted on campus.

As a result of the workshop, some basic building blocks were identified to compose the visual vocabulary of data protection and to originate more complex icons or pictograms. Such a compositional approach also derives by the underlying ontological modelling of the concepts. For instance, an arrow with gears means "processing", whereas "personal data" is represented by a folder with a user figure outline atop it. When personal data is processed, the basic personal data folder is combined with the arrow and a more graphically elaborated personal data folder to show the result of the processing activity (e.g. anonymized data). The visual narrative guides the reader from the left-hand side, where the basic personal data folder is showed, to the right-hand side, where the result of the processing activity is shown (e.g. Fig. 6.7). Thus, a standard way to combine the visual elements to achieve consistency across the icon set was developed. The need for coherence, precision and completeness across the icon set resulted in some complex icons, that could be defined more as pictograms and visual narratives rather than as single icons. This also derives from one of the tendencies that emerged during

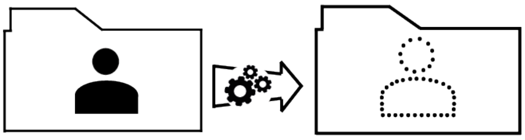


Figure 6.7: Example of icon composition for "anonymization": personal data (the folder on the left-hand side) are processed (the geared arrow) and result in anonymized data (the folder on the right-hand side).

the workshop: some groups tried to generate metaphors for some complex notions (e.g. for derived data), which were, however, read literally by users and not swiftly nor correctly interpreted. Thus, literal representations were chosen over their metaphorical counterparts.

This first icon set is composed of 18 icons (*see* Appendix C):

1. three icons describing different types of personal information, whose difference is especially relevant in the exercise of data subjects' rights: original, processed, and derived personal data;

2. five icons describing the main agents involved in data processing: data subject, data controller, data processor, third party, and supervisory authority;

3. eight icons describing processes carried out on the data: profiling, direct marketing[7], copying, automated decision-making, encryption, anonymization, pseudonymization, transfer to third countries;

4. two pictograms describing two data subjects' rights: right of access and right to data portability.

The evaluation of this icon subset is described in the next Chapter (*see* Sect. 6.4).

[7] Profiling and direct marketing can also be processing purposes.

6.2.2.2 The Design of the Second Icons' Subset

Two consecutive workshops were held at CIRSFID[8], University of Bologna, in March 2018. The first workshop [1] aimed to complete the icon set with the missing classes of concepts identified in the ontology and in Articles 13-14, whereas the second one [2] aimed to harmonize the design style among the two icon subsets. The vast majority of participants were legal experts and designers from the Academy of Arts in Bologna and the Academy of Arts in Florence.

The two workshops were structured around the design cycle explained above, over 8 hours, and could build on the strengths and weaknesses of the icon subset produced in the first workshop. Thus, the central elements of the visual vocabulary were provided (i.e. the data subject as a user, processing as gears, etc.) to be reused in the creation of the new icons. In order to create a coherent set of icons from the very beginning, three groups composed of a balanced mix of designers and legal experts were formed and to each group was assigned one of the following three classes of concepts: (a) data subjects' rights, (b) processing purposes, and (c) legal bases. Each group received a simplified definition and a practical example for each concept. Furthermore, the intended icons' context of use was specified: an exemplifying privacy policy with a structured layout was distributed (based on the layout shown in Appendix A).

Providing concepts organized in classes and providing the layout where icons would be inserted served to generate a coherent set. For instance, the very abstract concept of "processing purpose" must be seen in a global view where arrows exit a personal data folder and move towards a specific purpose (*see* Fig. 6.8). In addition, this complementarity of elements is a more usable and simple of evolution of the modular composition proposed in the first workshop and illustrated in Fig 6.7. Finally, specific instructions to generate ideas and sketch them out on limited space were given to the workshop's

[8] Interdepartmental Centre for Research in the History, Philosophy, and Sociology of Law and in Computer Science and Law `http://www.cirsfid.unibo.it/`.

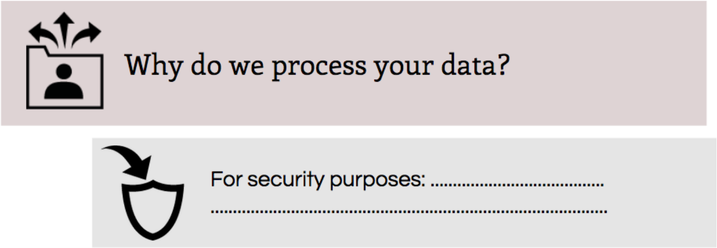

Why do we process your data?
For security purposes: ....
....

Figure 6.8: Example of icons' systematic use: the arrow exiting the personal folder stands for the general concept of "processing purposes" and heads towards a specific purpose, here e.g. "security purposes".

participants so that icons would be simple and ready to be displayed in small dimensions, such as on mobile devices.

Attention to balance among simplicity of representation, distinctiveness of some traits, but also coherency of elements across icons was also distributed among the guidelines. For instance, a hand's palm facing up was chosen to indicate any data subject's right. The metaphor underlying the holding hand represents the concept of being in control and having the power over the element located above it. The palm recurs in every data subject's right as common denominator among concepts belonging to the same class, but the meaning of the icon as a whole is specified by the element held by the hand: e.g. a bin for the right to erasure, a pencil for the right to rectification, etc (*see* Fig. 6.9). The hand element needed to be sufficiently big to be noticed as common denominator across icons, but at the same time sufficiently small to make individuals focus their attention on the distinctive elements placed on it in order to avoid similarity that neutralizes icon's distinctiveness.

The tension among icons' usability and their supposed informative value, which had prompted criticism during the evaluation of the first icon set (*see* Sec. 6.4), re-emerged persistently during this second workshop. On the one hand, legal experts warned of the risk of misrepresentation or oversimplification when data protection concepts. were visualized. On the other hand, exemplifying individuals representing an entire class and icons containing few

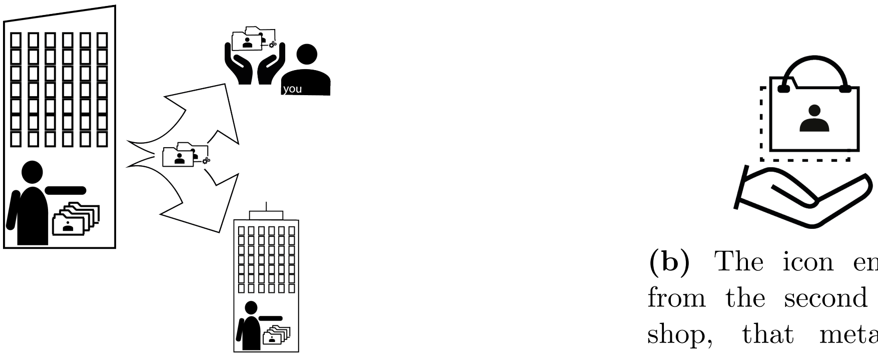
(a) The icon for the right to erasure

(b) The icon for the right to rectification

Figure 6.9: The icons present a common element, the hand, which stands for the class of subject's rights.

details must be preferred to achieve ease of recognition and adaptability to different contexts. For example, the comments received by the representation of the data controller in the first user study provoked a long and heated discussion. In the previous workshop and after some small sample testing, it was decided to represent the data controller as a person inside a building to convey the idea of a person in charge, inside of an organization. Following the first user study's results and the fact that the controller is a basic element that is combined with others to compose more complex icons (i.e. the legitimate interest of the controller, the contract, etc.), a simpler, even if arguably imprecise, exemplification of the controller was chosen: that of a person dressed as a business man.

Given the previous research on graphical symbols outlined in Section 6.1, it did not come as a surprise that concepts with a higher degree of concreteness (e.g. "contract"), for which an exemplification could be easily provided (e.g. "research purposes"), and that could rely on established visual convention (e.g. the bin to signify erasure in the "right to erasure" icon) were more quickly and effortlessly visualized. Conversely, abstract or general concepts such as "rights", "processing purposes", "service enhancement" and "service provision" were object of thoughtful consideration and, even, intense debate. Decisions that appeared arbitrary to some group members had to be taken. To cope with these difficulties, an assumption that had emerged in the previous workshop was expressly challenged, i.e. the fact that metaphors must be avoided to enhance clarity and reduce the openness of interpretation. Indeed, literal pictograms produced in the previous workshop (e.g. the right to data portability), despite their concreteness and their arguable informational value, cannot work in small dimensions and were considered too complex during the evaluation phase. On the contrary, a metaphor in which one idea is understood in terms of another is well suited to convey meaning through minimal elements. For these reasons, for instance, a folder in the shape of a suitcase was proposed to more compactly recall the right to data portability (*see* Fig. 6.10). Thus, many explanatory details of the first icon were lost but it was agreed that, if icons need to be usable and scalable elements, some specifications must be left to the written privacy terms that they complement. Conversely, if the visual elements aim to fulfil an explicative function, then different visualizations, such as pictograms, illustrations, and even animations, can be proposed (see Section 6.8.2 for a discussion).



(a) The icon that emerged from the first workshop and that was designed to depict the right to data portability in a literal manner.

(b) The icon emerged from the second workshop, that metaphorically depicts the portability of personal data as a suitcase with wheels.

Figure 6.10: The two icons realized for the right to data portability.

Furthermore, metaphors are used consistently throughout the entire icon set to convey meanings, i.e. a data folder to indicate “personal data”; an arrow leaving a circle of stars to signify “transfer outside the EU”; binary code to express something that is not readable by humans, thus “encrypted data”; gears to represent a functioning machine and, hence, “data processing”, although it is not a mechanical processing, etc. At different extents, they all are metaphorical depictions of a concept, but some are more familiar than others.

Iconographical choices, especially if metaphorical, were discussed at length in the individual groups and then in plenary, when harmonization with the visual elements generated by the other groups was also sought. Eventually, some of the icons produced during the first workshop in need of refinement (i.e. supervisory authority, controller, third party, right of access, right to data portability, etc.) were also presented and discussed, to gather feedback and alternative ideas concerning more functioning solutions. Comments, doubts, and ideas for promising visual solutions were recorded and later transcribed in a workshop report that was distributed to the participants some days afterwards in view of the subsequent workshop that aimed at the harmonization of the style of the two icon subsets and at their digitalization.

During this third workshop, a grid composed of squares of 16x16 mm (64x64 px ca.) was provided to transform the draft icons into digital form, by following the privacy policy template provided. Some visual solutions that worked on paper were not depictable in a smaller digital form, thus some icons' details had to be discarded. Providing coherence across the dimensions of traits and the elements of the set was also set as a priority.

At the end of the workshop, the previous data protection icon subset had been enriched by the following icons:

- 11 icons representing data subjects' rights: data subject's rights (as superclass), right of access, right to data portability, right of rectification, right of erasure, right to be informed, right to withdraw consent, right to lodge a complaint to the supervisory authority, right to restrict processing and right to object to processing (for this concept two different alternatives were produced because it was impossible to elect one best alternative internally);

- 7 icons representing legal bases for processing: legal basis (as superclass), consent, contract, legitimate interest of the controller, public interest, vital interest, legal obligation;

- 6 icons representing some common processing purposes[9] purposes (as superclass), statistical purposes, research purposes, security purposes, purposes of service provision, purposes of service enhancement.

These icons are displayed in Appendix D. An English translation of the simplified definitions provided to the workshops' participants is also shown, together with the reasons behind each iconographical choice. Also the first icon subset was re-elaborated following the visual conventions and the elements' dimensions established during the day. The evaluation of this icon set is illustrated in Sect. 6.5.

6.2.2.3 The Third Iteration Design

The evaluation carried out on the second icons' subset highlighted some critical points of the icon set in terms of legibility and comprehensibility. On the basis of the study results, a further, final elaboration of the icons was carried out in July 2018 at the CIRSFID, University of Bologna. This last workshop had three main goals: firstly, to redesign those elements of the icons that resulted less legible and recognizable (e.g. legal basis, right to data portability); secondly, to elaborate alternatives to those icons that had scored worst (e.g. purpose of service provision, legal obligation); thirdly, to harmonize every element across the icon set, especially in terms of their dimension and the line's thickness, and to simplify those icons that were yet too complex and detailed (e.g. the supervisory authority, automated decision-making). Moreover, three alternative icons for the data sharing with third parties, which is a fundamental concept that is omnipresent, especially in consent requests, were designed.

[9] These are the basic, recurring processing purposes identified in our analysis of the GDPR and inserted in its ontological formalization (together with the less frequent judicial, humanitarian, health-related, and journalistic purposes). However, service providers usually list many additional and more precise purposes in their privacy policies to justify the processing operations they carry out on data subjects' data. A comprehensive analysis of these could be carried out to discover if they all are individuals that can be attributed to one of these few ontological classes.

6.3 Evaluation

A two step-based approach for the evaluation of DaPIS was developed and illustrated in [223], given the characteristics and functions of the data protection icons. Firstly, icons must be evaluated as stand-alone elements, i.e. according to dimensions such as comprehensibility and legibility. In the second place, icons must also be evaluated for their function in context, as information markers that support the navigation through large amounts of information and increase speed and accuracy of comprehension.

In the following, the first stage of evaluation is described, i.e. the evaluation of icons as stand-alone elements. The first section (sect. 6.3.1) provides an exposition of classical assessment measures, followed by the description of three subsequent user studies that were carried out to gauge the effectiveness of DaPIS (Sect. 6.4 for the first study, 6.5 for the second one and 6.6 for the third).

6.3.1 Evaluation Measures for Graphical Symbols

Given the prominence of graphical symbols e.g. on graphical user interfaces (GUIs) or in public spaces, there exists a body of literature concerning methods and relative measures to assess the effectiveness of symbols along different dimensions.

6.3.1.1 Ease of Understanding

Ease of understanding is “the most important single index of a symbol’s effectiveness” [84, p. 292]. A typical measure of evaluation is hit rate, i.e. the number of correct matchings between an icon and its referent. The only international standardized existing methodology for the comprehension of graphical symbols (ISO 9186-1.2014 [167]) is unsuitable because it has been designed for symbols meant to be employed in public spaces (e.g. airport) and whose referent, i.e. the entity to which the symbol refer, is known to users (e.g. airplane). ISO 9186-3.2014 is aimed at assessing the ease of association between icon and referent, after familiarity training for unknown referents has been carried out. However, the process of learning the referents' meanings takes time, which increases with big number of referents, as in the present case. Therefore, such a test will be taken into account for the future, when familiarity will be rehearsed and an appropriate experience to increase motivation to learn will be designed (*see* also Section 7.1). As for what concerns the ETSI Multiple Index Approach [163], it was meant to evaluate symbols for telecommunication interfaces of the early 1990s, and to elect the best alternative among several icons for the same referent.

As recalled earlier (*see* Sect. 6.1.2), DaPIS has some distinctive qualities that make such standardized evaluation methods ineligible for this context. As a matter of fact, individuals are usually not familiar with the referents (e.g. the concept of 'pseudonymization'), whereas the icons might entail low concreteness and high semantic distance. Moreover, many icons are only marginally based on a shared visual vocabulary, thus familiarity with some graphical conventions is expected to be low or even non-existent. It is hard, then, to reach high rates of comprehensibility at first exposures and to set a high bar for acceptance. However, the provision of sufficient contextual information about where the icons might be found and about their function in context should lead to better results, as shown in different contexts [299]. Nevertheless, before testing the icons' functionality in context, icons must be evaluated as single elements to provide detailed insight into the mental models and the line of reasoning behind the interpretation process and to isolate the variables that ease or conversely worsen icon recognition.

6.3.1.2 Legibility

Legibility, namely the ease of recognition of the elements that compose an icon determines the ease of recognition of the icon as a whole. This dimension is important because if some elements are not easily visible (e.g. for their size) or recognizable (e.g. for the way they are designed), they could hinder the comprehensibility of the icon's meaning.

6.3.1.3 Subjective rates

Subjective certainty can be also considered to be included in the test because high uncertainty can reveal higher possibilities of incomprehension. Qualitative feedback should also be encouraged to understand the rationale behind certain preferences or rejections and even to increase comprehension scores [299].

6.3.1.4 Best alternatives

If there are multiple alternatives for the same concept, it is good practice to ask for a preference among them (*see* e.g. [163, 125]) In many cases, however, the icons of our icon set have been created for newly introduced concepts and there are no alternatives. Furthermore, some alternatives for the same referent have been already discussed and discarded during early stages of the design process. Nevertheless, some alternatives for those concepts that had not found a good visual representation had to be evaluated.

6.4 Evaluation of the First Icon Design

6.4.1 Introductory Considerations

The first user study (*see* Appendix F), carried out in August 2017 in the US, did not focus exclusively on the first icon subset (*see* Sect. 6.2.2.1), but also on the efficacy of other channels to convey the same data protection concepts: simplified definitions, real-world scenarios, and classical legal terms. For the scope of this research, however, only the data around the evaluation of the icons will be reported. The study was conducted through in-person observations and interviews with participants. The subjects had to perform different tasks, explained below, and follow a think aloud protocol. They were thus asked, throughout all the tasks, to verbally express their thinking process and the reasons behind their choices. When they went silent, they were encouraged to verbalize their thoughts.

6.4.2 Participants

The research study was arranged for 20 participants, from as diverse as possible demographics, in terms of origin, gender, age, education, and profession. Of the 16 participants that showed up, 7 males and 9 females, all of them described themselves as “having lived most of their life in the US”. Indeed, the participants were recruited in the Bay Area (San Francisco), where the study was carried out, and received a 30$ Amazon card gift as an incentive to take part in the study. Their educational background was diverse, but still medium-high: all the participants were at least high school graduates. Equally diverse was their profession, and nobody had a legal background. Their age also varied, with a minimum of 19 and a maximum of 76, with an average age of 41 and a median of 39.

6.4.3 Tasks

Task 1: Rephrase 18 simplified definitions. The test participants received a piece of paper with 18 simplified definitions describing the concepts of the 18 icons listed next to them. The order of the definitions and of the icons was randomized for each participant to prevent possible order effects. After a brief contextual introduction, the participants were asked to read and restate in their own words the definitions or mention examples explaining the definitions. This step had the goal of assessing users’ understanding of the definitions and to point out possible flaws in the wording. The underlying idea is that anyone should be able to understand the definitions, even without any previous knowledge of the topic. To ensure the definitions’ comprehensibility, from the lexical point of view, the 4000 most commonly used words according to the Collins dictionary were exclusively employed. Even the syntax and the lexicon were adjusted to reach the highest possible level of readability, according to measures such as the Flesch Reading Ease Score (FRES) and the Flesch-Kincaid Grade Level (FKGL), so that the definitions would be as understandable as possible (min. FRES = 59.6, max. FRES= 92.9; min. FKGL= 2.2, max. FKGL=8.4). The simplified definitions were double checked by the team's legal experts to ensure the correspondence of their meaning to the definitions in the Regulation.

Task 2: Match between definition and labelled icon. Then, the participants were asked to find the best icon match for each definition, in a one-to-one correspondence, among the 18 icons listed on the other side of the sheet of paper. However, the participants were also allowed to choose even more than one icon per definition, or viceversa, when they felt that there was no best match. They were encouraged to explain the reasons behind the icon choice. The icons reported in Appendix C were coupled with the correspondent labels taken from the ontology, since it is recommended to join the visual representation with a descriptive keyword to reduce chances of misinterpretation. It was deliberately decided to employ the legal terms used to describe the concept, in order to explore how easily these terms can be understood by average users.

Task 3: Post-study self-reported effort rating. Finally, the users were asked to rate the effort for each communication modality on a seven-point Likert scale, and to explain the reasons for a certain score. Room for further comments, questions, or suggestions was also allowed, especially concerning icons.

6.4.4 Analysis

6.4.4.1 Qualitative analysis

An analysis of the interviews' transcripts was carried out with a twofold purpose. Firstly, to gather the qualitative feedback on icons and other communication modalities, e.g. common types of problems or patterns, and, secondly, to determine the subjects' level of understanding of the simplified definitions. Whereas usual icon recognition tests employ familiar referents that participants need to associate with an icon, in this case many referents are specific of European data protection law, thus it is reasonable to expect that they will be unknown to the test participants. It would have been, therefore, very difficult to determine whether a wrong association between referent and icon would have depended on the characteristics of the visual element or on the lack of familiarity with the concept itself. Thus, the rephrased sentences for Task 1 were extracted from each interview, compared, and rated.

6.4.4.2 Measures

In this first user study, classical measures for icon evaluation were reproduced.

For Task 1 (rephrasing the simplified definition):

Answer accuracy: measure that considers the correctness of each answer. For each simplified definition, the accuracy was calculated as the sum of the scores accrued from each participant. The measure ranges between 0 (no participant could understand the simplified definition or the scenario) and 16 (all participants could understand the simplified definition or the scenario) for each item. Each answer could receive a score of 0 (in case of wrong, irrelevant or unknown answer), 0.5 (in case of vague, incomplete or partially wrong answer), or 1 (in case of precise, correct answer or relevant example). The maximum score was assigned only when the subjects restated the sentence clearly, precisely, specifically, correctly in their own words, provided a correct example, or mentioned the concept to which the original sentence refers. For example, for the term ‘data subject’, with the original sentence: “this is the person to whom personal data refer”, a maximum score was assigned to the participant (P9) that restated: “That would be me, because it’s my data so it’s referring to me”. A low score indicates that the definition was difficult to understand.

For Task 2 (match between definition and correspondent icon):

Hit rate: (correct matches between referent and icon + partial correct matches *0.5) / total number of given matches. Partial correct matches are those cases where there was a wrong match alongside a right match so only half point was computed. A low score reflects the fact that the association was difficult to make.

Error rate: number of wrong matches between referent and icon / number of total given matches. The closer to zero, the lower the number of errors, thus the more understandable the simplified definition and its connection to the icon.

Missing values: number of lacking matches between referent and icon / number of total actual matches. This measure reflects the level of incertitude about the correspondence between definition and icon.

For Task 3 (post-study self-reported effort rating):

User experience measure: self-reported rating expressed on a scale with values ranging between 1 and 7, done after the other tasks. The user experience of the different communication modalities (icons, simplified definitions, scenarios, legal terms) was evaluated in terms of self-reported effort rate.

6.4.5 Results

6.4.5.1 Task 1: Understandability of simplified definitions

Overall the understanding average of the definitions was 52% (min= 25% and max=78%). Exactly half of the definitions are above 50%, while the other half below. In Fig. 6.11 are reported the single rates of understanding for each simplified definition.

‘Direct marketing’ was the better understood definition by the participants. Throughout all the tasks, indeed, most of the subjects made continuous reference to online and offline direct marketing practices, one reason probably being that they have firsthand experience of it. definition Right to data portability' ranked second: this result is somehow unexpected since it is a newly introduced right by the GDPR. However, portability as transfer of information from one entity to a new one can be a familiar concept, as the examples provided by the subjects reveal: the transfer of data to a new device or the transfer of medical records to a new practitioner. Then followed 'data subject', 'supervisory authority', and 'profiling'.

Conversely, the definitions of 'processor', 'third party', 'processed personal data', and 'pseudonymization' ranked last. This can either be due to the lack of clarity of the definition itself or to the understandability of the concept described. Since the participants were only given a minimal context at this stage of the test, the results are not dissatisfactory on the whole. Indeed, some subjects found the definitions rather vague or commented that more context was needed to understand what the definitions referred to (*see* 6.4.6).

6.4.5.2 Task 2: Match between simplified definitions and icons

The average success rate of the matching between icons and simplified definitions was 69% on the whole, whereas the success rate for each icon is displayed on the graph in Fig. 6.12. 'Copying' was recognized 100% times, closely followed by transfer to third countries, and direct marketing. The matches that scored worst are 'processor' and 'controller' (but *see* conceptual confusion between the two in Section 6.4.6), 'derived personal' data and 'pseudonymization'. The latter also received the highest number of missing values, therefore indicating high incertitude (*see* Section 6.4.6).

As can be noticed from the graph in Fig. 6.13, in almost the totality of cases the subjects performed better in the matching task than in the rephrasing task. The reasons can be several. The simplest one is that, by having tried to rephrase and thus to reflect upon the definitions beforehand, the participants had already built a mental model of what the definitions referred to, so it was easier to associate the definition to the corresponding

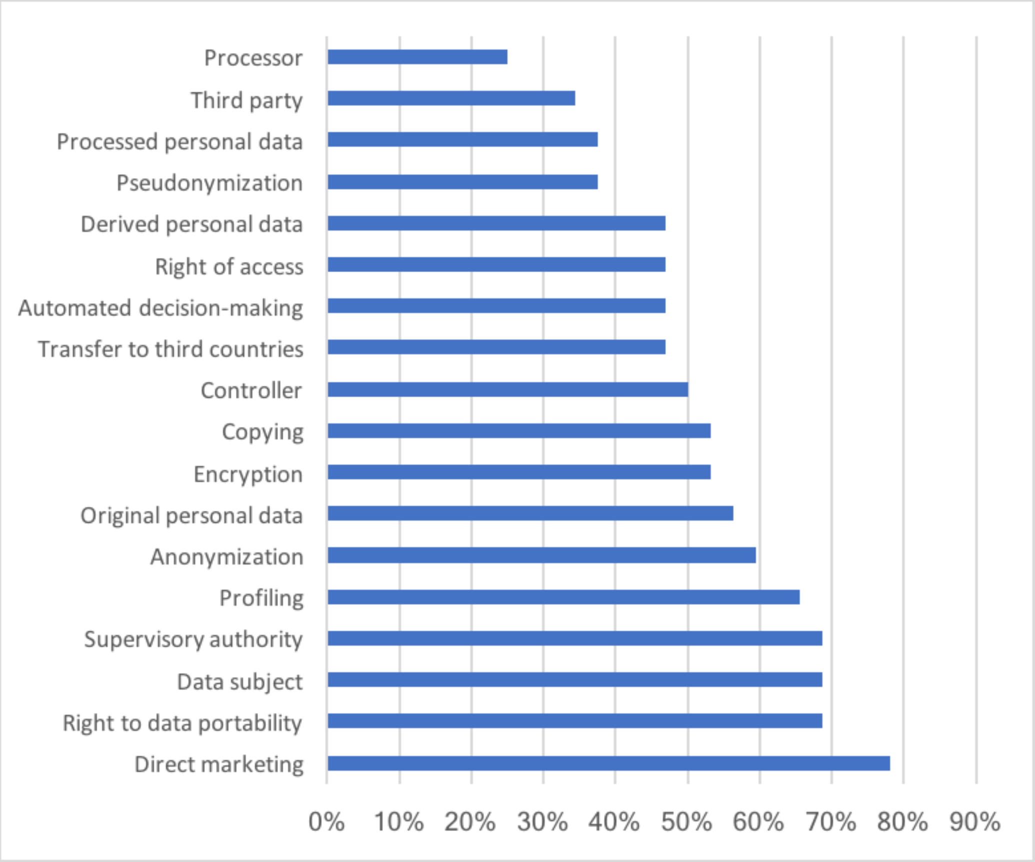


Figure 6.11: Chart that represents the understandibility relative to each simplified definition (Task 1), where the closer to 0 the less correct rephrasing

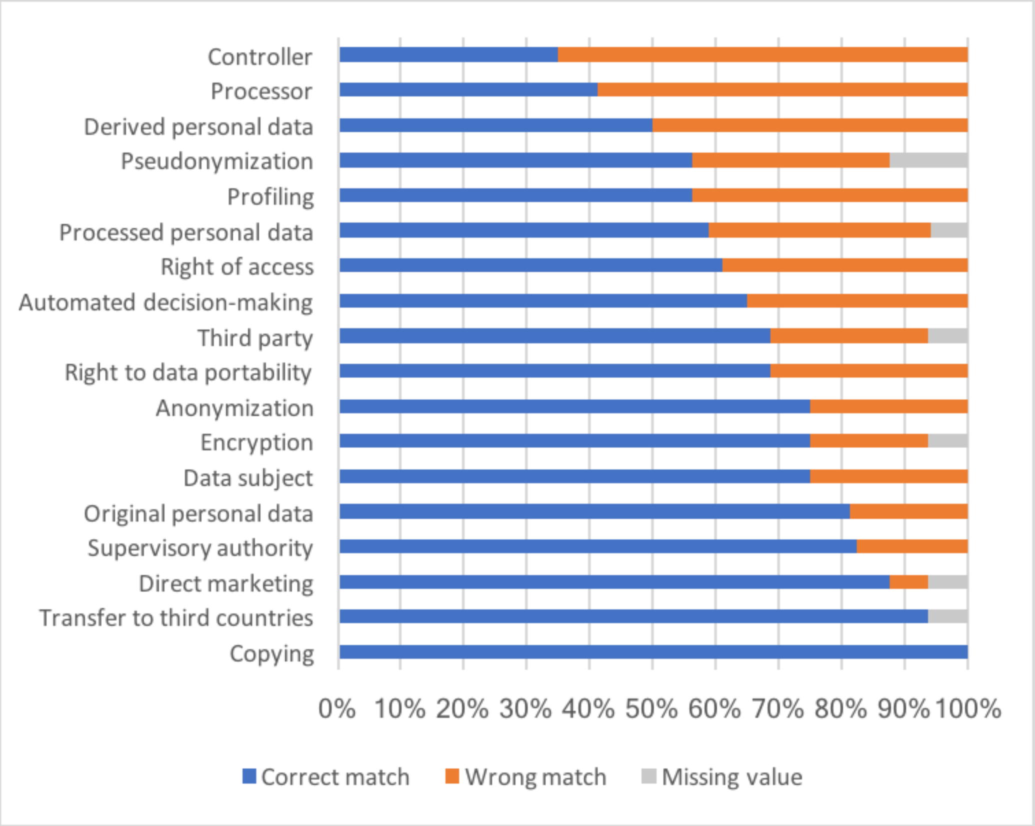


Figure 6.12: Chart that represents the correct matches, the wrong matches, and missing values for the association between icons and simplified definitions (Task 2)

labelled icon. However, it was also observed that the participants used some visual elements of the icons to determine the meaning of the labels or to perform the match with the definition. For instance, in the case of transfer of personal data to third country, whereas the results from the rephrasing task highlighted that the definition must be improved, the presence of an arrow exiting a circle of stars made it clear that there was ‘movement of personal data outside of the European Union’.

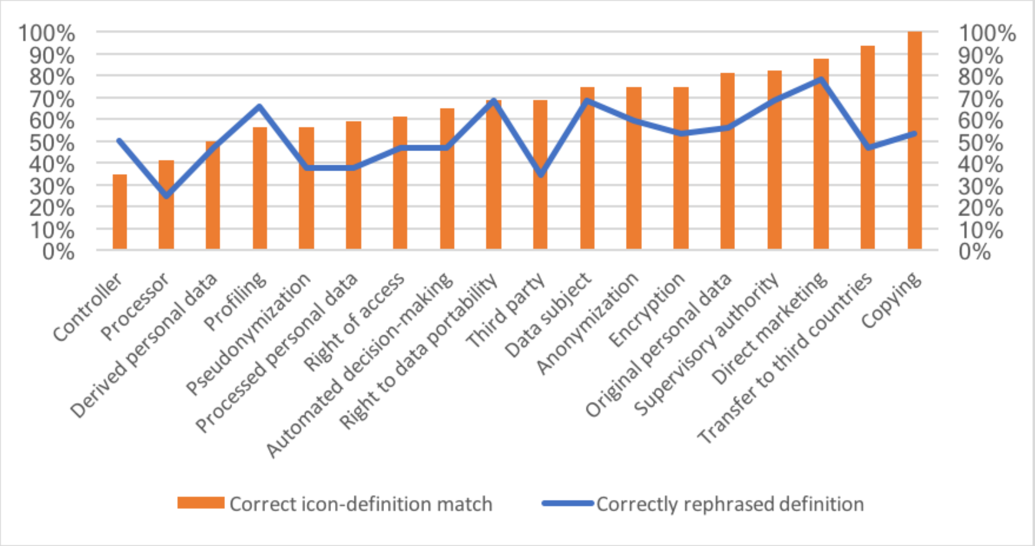


Figure 6.13: Chart that compares the correct rephrasing of the simplified definitions (Task 1) and the correct matches between icons and definitions (Task 2)

6.4.5.3 Task 3: Subjective evaluation of effort

A Friedman test on the overall effort ratings returned a significant result (2= 12.867, p=0.004933), meaning that there is a detectable difference in the effort ratings that the subjects gave for the four different communication forms (i.e. icons, simplified definitions, scenarios, and legal terms). A Wilcoxon signed-rank test (adjusted with Holm's sequential Bonferroni correction) revealed a statistically significant pairwise difference (p=0.0135827) between the ratings of icons and scenarios. In general, as also the boxplot (Fig. 6.14) shows, our icons were perceived to be more difficult to understand than the other communication modalities (*see* Sect. 6.4.6).

6.4.6 Discussion of Results

The self-reported evaluation draws attention to the fact that there were contrasting opinions about DaPIS. From a thorough analysis of the interviews' transcripts, common patterns and recurring comments on the icon set were gathered.

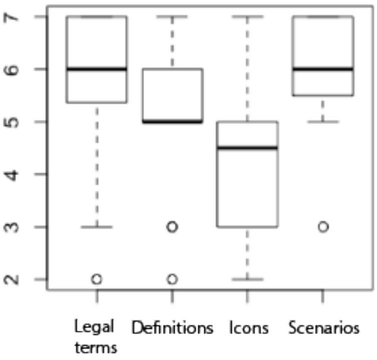


Figure 6.14: Boxplot graphs that represent the effort rate on the different communication modalities. On the vertical axis is displayed the effort rate on a scale from 1 to 7 (Task 3)

Low Familiarity with the Icons' Referents: The analysis revealed that one of the reasons of recognition or non-recognition of the icons was the familiarity with their corresponding referent. For instance, the subjects named many examples of marketing practices throughout the whole test, referring to their own personal experience. Indeed, 'direct marketing' was among the more easily identified icons. This tendency is even more evident among the least recognized icons. 'Pseudonymization' constitutes perhaps the most emblematic case in this sense. Many subjects struggled to understand the concept underlying it. Unlike anonymization, the notion of pseudonymization is reasonably unknown to the general public. Finally, many participants encountered problems even only when reading the label associated with the icon and some asked for explanation of its meaning. From this follows that, even if the icon is well designed, people notably rely on the associated label to interpret the visual element. Thus, even the label must be carefully conceived with attention to user-friendliness in order to foster understanding, instead of resorting to the legal terms.

Shared Visual Vocabulary: It was also observed that the participants relied on some icons' features to determine their meaning, especially when familiarity with the visualization was high. For instance, in the case of ‘transfer of personal data to third country’, the presence of an arrow exiting a circle of stars was correctly interpreted as movement of personal data outside of the European Union. The use of the file folder for ‘personal data’ and of the user outline for ‘data subject’ were positively evaluated, since they rely on a graphical language that is already part of the shared knowledge of computers’ and social networks’ users. These results confirmed that the more the icons rely on an established visual language and common mental models, the more recognizable they are. Some elements of the shared visual language on data protection have been successfully identified, but more research towards this goal is needed. However, if it is the concept to be unknown, it is difficult to find a sufficiently good visual representation that can straightforwardly communicate its meaning.

Distinguishing Features: ‘Processor’ was frequently confused with ‘controller’, and vice versa. This is due not only to the fact that these two roles were not readily distinguished at a conceptual level, but also to their visual representations. Both processor and controller are depicted as user figures inside a building, the former overlooking file folders, whilst the latter overlooking processing gears. The participants’ comments highlighted that the difference between the two icons was perceived as too subtle, thus went in some cases unnoticed. From this follows that consistency of the elements across the icons is important, but greater relevance must be given to the distinguishing features of the icons. Similarly, the distinction between the different personal data depicted in the rights was ignored. Thus, a visible, straightforward way to show the distinction among similar concepts must be envisioned.

Chances of Misinterpretation: From this user study, one risk concerning misinterpretations emerged: ‘profiling’ was repeatedly mistaken for ‘direct marketing’, and vice versa. On the one hand, this might indicate that individuals are aware that the two processes are closely interlinked. However, this poses a serious problem when consent is asked, since individuals can choose to give consent to none, either one, or both operations. Being able to distinguish between them not only at the conceptual level, but also at the visual level is therefore crucial, if consent was to be asked through clickable icons [31]. Also 'automated decision-making' exposed a similar problem: although the computer depicted in the icon helped some subjects to correctly match it to the concept, it was sometimes confused with profiling and was more generally associated to artificial intelligence. Although such exchange might indicate a non-trivial understanding of this technical notion, from a legal perspective it might also cause some issues: automated decision-making refers solely to those decisions that have significant effects on data subjects, such as the eligibility for money loans. The icon was therefore not able to be as precise as the legal experts inteded to clearly represent this distinction. This may entail considerable consequences, because it is connected to the possibility of exercising the right to object to an automated decision.

Combination of Icon and Label: In general, however, not only the users relied on the combination of icon+label for the interpretation, but in many cases they principally relied more on the textual cue than on the visual element. Half of the participants even mentioned the fact that without labels or some extra information (e.g. a clickable, pop-up definition), the images would be hard to understand or ambiguous. Two participants (P1, P6) even pointed out that, if users are not familiar with the icon's referent, the visual is not going to help. Thus, a user-friendly label must always be associated to the visual representation of a data protection concept, especially for first exposures to less semantically transparent icons. Thus, the provision of both textual and visual cues in a solution should be preferred, so that individuals can leverage either one or both of them to understand the communication.

Non-Native Speakers: Nevertheless, there are cases where this is not true: one participant (P8), whose native language was not English, relied strongly on the visual elements to match it with the definition, especially when she did not know the meaning of the label associated to the icon (e.g. 'encryption', 'anonymization', 'pseudonymization', etc.). And in those cases, her assumptions were in fact correct. If it was possible to give an easily intelligible visual representation of complex data protection concepts, it could arguably benefit individuals with lower literacy levels.

Simplicity Versus Precision: Probably the most important insight derived from this study is that some of the icons were deemed "too complicated" and "too crowded", such as the pictogram representing the right to data portability (*see* Fig. 6.10). However, the level of detailedness of this representation helped some of them to follow the embedded narrative and, thereby, understand how this right unfolds and distinguish it from the right of access, which presents similar elements. A trade-off between accuracy and coherence of representation, necessary for legal purposes, and simplicity, as users require, clearly emerged as future direction of research and was the main concrete guideline the guided the icon redesign in the following design worshops.

6.4.7 Limitations of the study

The study described in the previous pages present some limitations. For instance, a focus was placed on the users' verbal skills in the rephrasing tasks, whereas in the future it would be important to discover the effects of other learning styles (i.e. the preferred way for individuals to process and remember information) on the understanding of different communication modalities.

It must also be acknowledged that the match between icons and definitions was finite. Therefore, as some participants even pointed out, they carried out the task through an elimination process, instead of abstaining from those matches they were most uncertain about. However, in the real world the set of referents for a certain icon or a certain definition is not closed, it is rather infinite.

Moreover, the subjects sample was not very diverse in terms of educational background and nationality. It would be therefore necessary to carry out a user testing with people with lower education levels and different nationalities, especially Europeans. The logistics of the study, physically based in the US, caused the sample to be almost exclusively American. Nevertheless, there is value in US participants' feedback to DaPIS, since it aims to become a standard set of icons that is understandable across nationalities, language, and cultural backgrounds. The goal is to reconcile different testers' feedback to the icons to create a set that can by itself communicate effectively to multiple audiences. Alternatively, icons that adjust depending on the audience might be recommended, as it will be proposed in the last chapter.

Finally, some more limitations are discussed in the next section, since they guided the design of the evaluation phase for the second subset of DaPIS.

6.5 Evaluation of the Second Icon Design

6.5.1 Introductory Considerations

The evaluation of the second subset (*see* Sect. 6.2.2.2) was carried out across different dimensions compared to the first user study, in order to avoid its limitations. Firstly, the focus was solely on icons to gather as much feedback as possible, whilst other communicative devices were ignored. Relevance was given to three aspects: icons' legibility, correspondence between icon and it underlying concept, and alignment between users' and designers' mental models.

Legibility of the icons was examined to ensure that all the elements could be easily visible and recognizable even at small dimensions, since low legibility levels can influence recognition and interpretation.

A second dimension was ease of understanding of the icons. As recalled at the beginning of this Chapter, specific characteristics of DaPIS render it hard to replicate existent standard evaluation methodologies that employ quantitative measures such as hit rate. Such a measure was used to evaluate the first icon subset, but showed limitations since it was sometimes hard to determine whether wrong associations depended on the icon or rather on the individuals' limited knowledge about the underlying concept. The rephrasing task was adopted as solution, but the concrete test revealed that in certain cases it was arduous to establish whether low comprehension rates were due to flaws in the definition itself or to difficulties in grasping the underlying concept (*see* Sect. 6.4.5.1). Although the definitions were translated into simple language, the strict application of readability estimation in some cases transformed the descriptions into texts that were too simplistic to be easily understood.

For these reasons, a different approach was adopted in the user study described in the following pages: the same simple definitions provided to the designers during the icons' creation were displayed next to the icons and, instead of measuring efficiency of association between concept and icon, the primary focus was placed on the process of interpretation of the visuals. The adoption of such a method had three main goals: firstly, to clarify the difference between poor rates caused by icons' representation, and poor rates derived by lack of understanding of the underlying concept. Secondly, no interpretation strategy based on the exclusion of previous associations was deployed since this strategy could hide the level of subjective certainty about an association: in the previous user study, the icons selected and associated first were in most cases those that the data subjects could more easily recognize (e.g. copying). Finally, this strategy allowed to explore whether the rationales behind the iconographical choices made during the design phase could be grasped, i.e. if alignment between designers' and users' mental models was possible, even on less semantically transparent icons.

Moreover, unlike the first test, the icons were not associated with the corresponding label because it was found that subjects would base their interpretation most prominently on the latter than on the visual cues. Nonetheless, a label was placed above the definitions to which the icon corresponded, since it represented a contextual aid as support for the interpretation.

An additional consideration that drove the design of this second user test sparked from the fact that, especially for abstract concepts, any visual representation could be potentially suitable. After repeated exposure, individuals will eventually learn and memorize the association between an icon and its referent. It is for this reason that it was deemed necessary to assess the capacity of an icon to convey its meaning in isolation to replicate the conditions of those that will briefly glance at the icons, instead of reading the privacy policy's terms. Research on the effectiveness of contextual elements (e.g. labels, definitions, text, etc.) at complementing, reinforcing or even guiding the interpretation process is left for the future (*see* Section 7.1).

6.5.2 Participants

16 participants, 11 females and 5 males, were recruited in Bologna, where the study took place, through paper ads and online ads, and received a 20 euros Amazon gift's card as reimbursement for their time. The participants' age span ranged from 20 to 29 years old and the great majority were university students, more than two thirds with a Bachelor's degree, indicating a high educational background. A self-assessment of the participants' digital and legal skills was also asked. Three quarters (n=12) of the subjects described themselves as having intermediate digital skills, while two fell on both side. Half of the participants (n=8) also claimed to have basic legal competences, while two declared them to be non-existent and 6 intermediate. All of them had native or comparable levels of Italian: this was a necessary prerequisite to carry out a task where also comprehension of legal written definitions was requested.

6.5.3 Tasks

The study was conducted as in-person observations and interviews by three researchers at the University of Bologna, Italy, in March 2018. The test was constituted by a predefined set of questions (*see* Appendix G) along the dimensions explained above. The subjects were asked to answer such questions and were actively encouraged to follow a think aloud protocol. This test was carried out in Italian: in these sections, questions and answers have been translated into English. Researchers intervened when the users asked for explanations or examples around a certain concept, since the understanding of the concepts' meaning was critical for the experiment (*see* earlier at Sect. 6.5.1). To provide contextual elements to help the participants to create a mental model similar to the actual icons' context of use, a brief explanation about the research was given at the beginning of the test and a mock-up of a visualized privacy policy was shown. The participants were asked to record their answer in written form next to each question. The researchers took notes of participants' behaviours or of comments that were not recorded by the participants themselves.

Four tasks were designed:

Task 1: Icons' legibility: legibility was estimated by asking the test participants to name the elements composing each icon. Even the icons from the first subset were evaluated in this sense, since their style was harmonized during the third workshop and they had not received any legibility evaluation in the previous test. To replicate non optimal conditions (i.e. the worst case scenario), the icons' legibility at reduced sizes was explored. Hence, they were printed out at 16x16 mm to reproduce small settings and low resolution, as they could appear on paper-based privacy policies or devices' boxes. Furthermore, research shows that human beings make sense of information differently on screen than on paper (e.g. [195]), even if contrasting results exist (e.g. [250]), although these studies are rather focused on textual information than on visual information.

Task 2: Subjective rating on the icons' correspondence with its underlying meaning: for the reasons outlined above (Sect. 6.5.1), a subjective rating on a Likert scale that ranged from 1 to 5 (where 1=strongly agree, 3=don't know, 5=strongly disagree) about the ability of a certain icon to represent the corresponding concept, expressed through a label and a simplified definition, was preferred to a match task. Explicit explanations for the mark were asked and recorded in written form.

Task 3: Alignment between users' mental models and designers' mental models: the participants were asked to attempt to provide an explanation for the visual choices made by the designers. This task was meant to find out whether users could understand the reasons behind the choices and thus align their line of reasoning with the designers', despite their opinion on the appropriateness of a certain icon for a certain concept, evaluated in the previous task. In order to avoid influence by different wordings, the same identical simplified definitions that were distributed to the designers to spark the design process, were also provided to the test participants. Neither examples nor further explanations were given with the hand-outs, but rather provided to the subjects orally by the researchers if needed, similarly to the design phase.

Task 4: Best Alternative: one single alternative choice between two icons representing the concept of 'right to object to processing' was asked.

6.5.4 Analysis

The data collected in written form by the participants was gathered and integrated with the notes taken by the researchers during the study. This data was then analyzed by one of the interviewer in search for common patterns. Given the nature of the tasks and the study goals, qualitative analysis was the main source of data.

For legibility (task 1):

Hit rate : For each icon, there was a lower bound equal to a minimal set of elements that had to be identified in the icon for the answer to be considered correct. Wrong identification of elements and common mistakes were recorded.

For the correspondence between icon and concept (task 2):

Average : mathematical mean across the users' self-reported marks was computed for each icon;

Frequency : since means can hide details, frequencies for each mark were computed instead.

Furthermore, the explanations provided by the participants to motivate their choices were analyzed to find common lines of reasoning, but also to find out explicitly about the words that were employed to describe why some associations were easier or harder than others.

For the alignment between users' and designers' mental models (task 3):

Hit rate : number of correct explanations provided by the participants with respect to the designers' reason behind a certain iconographical choice, for each icon's element (e.g. 'right to erasure': one mark for the hand and a different mark for the bin symbol). Since some elements appear in more than one icon (e.g. the hand in all data subject's rights), the score for each element is computed as: number of correct matches/(number of responses given * number of icons where the element appears). Higher scores correspond to better user's understanding of the designers' reasons for a certain iconographical choice.

Error rate : number of wrong explanations provided by the participants with respect to the actual designers' reason behind a certain iconographical choice for each icon's element. The overall score for each element is computed as number of wrong matches/(number of responses given * number of icons where the element appears). Higher scores correspond to worse user's understanding of the designers' reasons for a certain iconographical choice.

For these last two measures, marks were assigned as following: + 1 if the user's explanation coincided with the designer's intention; -1 in case of wrong explanations; 0 when no explanation is provided, or the explanation is incomprehensible or vague (i.e. "because it's clear, well-known, intuitive" etc.), or the participant admits that she does not know the reasons behind a certain visual choice.

For the best alternative between the two icons for the "right to object to processing" (task 4):

Higher number of preferences : the best alternative was chosen simply by counting which of the two icons got the majority of votes. Reasons for the preferred choice were also recorded.

6.5.5 Results

6.5.5.1 Task 1: Legibility

Positive results were achieved on almost all icons, meaning that the elements were simple enough to be easily recognized, even in small dimensions. This confirmed that the simplification of icon design brought favorable results. The only icons that show lower rates of legibility are:

1. right to data portability: only one fourth of the respondents identified the bag-shaped file folder, whereas the great majority did not recognize it; three interpreted the drawing as a padlock;

2. right to lodge a complaint to a supervisory authority: almost all test participants could not correctly identify the file folder below the supervisory authority and almost half of them interpreted the gears as a key;

3. controller: only two subjects expressly noticed that the silhouette has a white shirt and is slightly different that a usual user silhouette;

4. legal basis: more than half of the participants could not determine that the element under the hammer is a column;

5. vital interest: some minor doubts (three people) on the graph within the hands;

6. encryption: almost everybody could detect the written characters in white, but only half could safely assume that it was binary code;

Although it was not a goal encompassed by this first part of the test, some participants attempted to provide a free interpretation of the icon even in this phase and the investigators let them free to do so (*see* also Section 6.5.5).

6.5.5.2 Task 2: Rating on Fitness of Correspondence

The results of the assessment on the icons' capacity of representing the underlying concept are reported in Fig. 6.15. The results reported in the following paragraphs are organized in three groups, according to the average value obtained by each icon:

1. best rated icons: average value ranging between 1 (=completely agree with the fitness of correspondence between icon and concept) and 2 (=agree);

2. medium rated icons: average value ranging between 2 (=agree) and 3 (=uncertain);

3. worst rated icons: average value ranging between 3 (=uncertain) and 4 (=disagree).

Best rated icons: Among the icons that scored best (*see* Fig. 6.16), the symbol of a bin to signify "cancellation, erasure" **right to erasure**

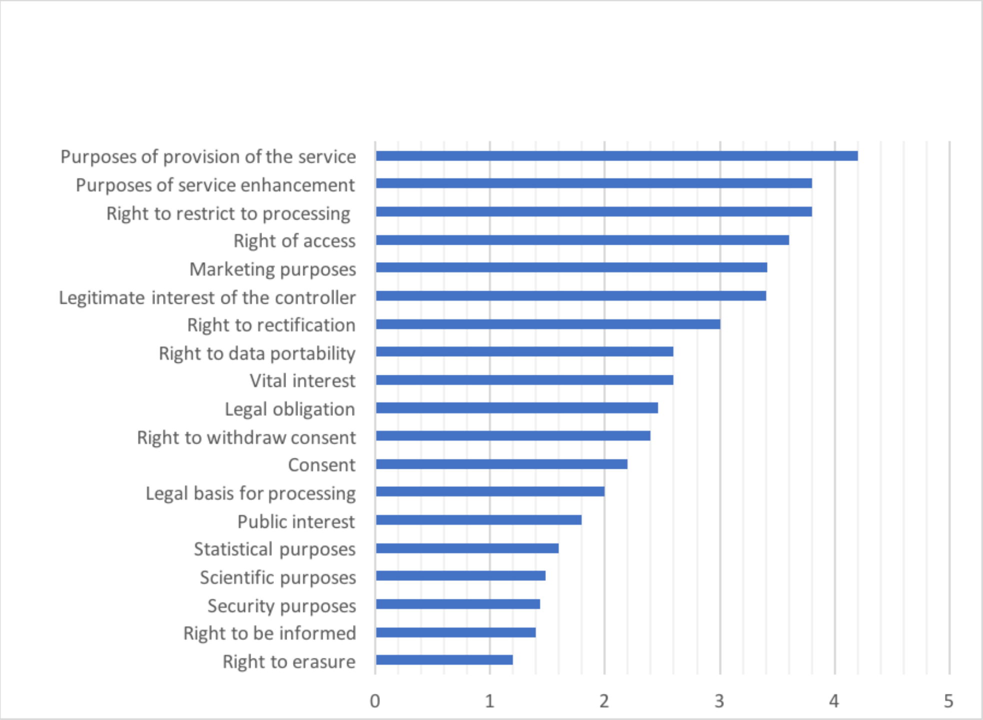
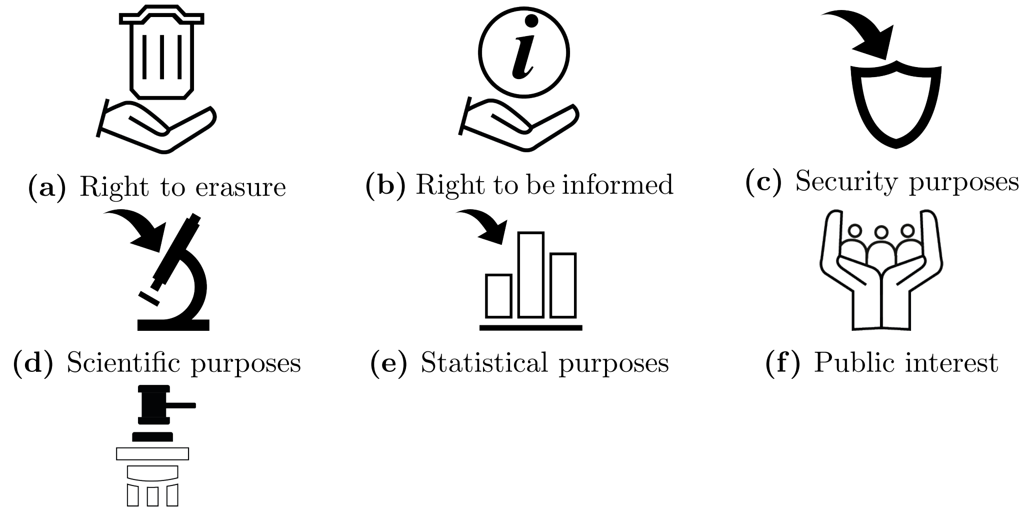


Figure 6.15: Means of the self-reported values for ease of understanding, where the closest to 1, the better the results

and of an “i” to signify “information” in **right to be informed** were described as “universal, immediate, instantly recognizable, clear, intuitive, unmistakable” because “grounded in our culture, codified and common on application software”.



(a) Right to erasure **(b)** Right to be informed **(c)** Security purposes

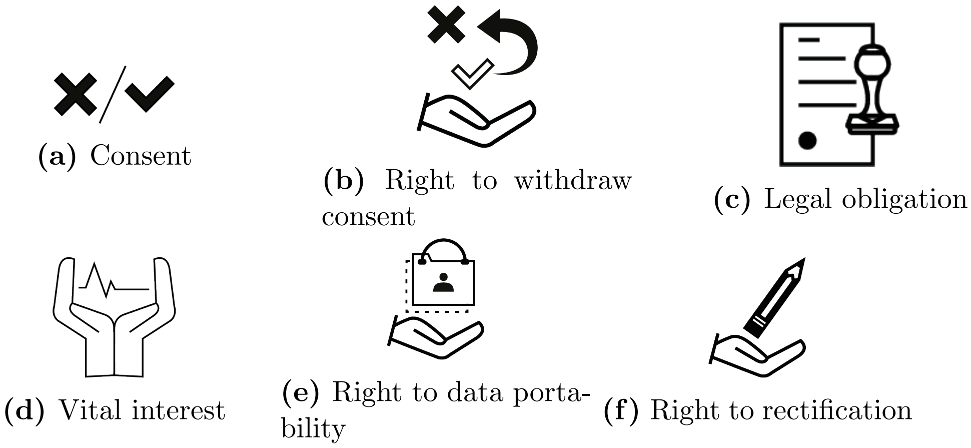
(d) Scientific purposes **(e)** Statistical purposes **(f)** Public interest

(g) Legal basis for processing

Figure 6.16: The best rated icons in Task 2

The **security purposes** and **research purposes** icons were also rated positively, since the shield is “stereotypical” for security, defense, and (antivirus) protection, whereas the microscope is “emblematic” of science and research. The bar graph “intuitively” recalls statistics (**statistical purposes**), whilst the presence of three user silhouettes seem to be easily associated to the idea of “public”, group of people or community (**public interest**) (for the meaning of the hands that sparked some doubts and lower scores *see* Section 6.5.6). In the **legal basis for processing**, only one respondent was able to link the column to the basis (but this might be caused by the low recognizability of this element), whereas the hammer was unequivocally associated to the legal sphere, which might be a rather vague association, or even more often to justice, which is technically incorrect. Nevertheless, the legal and juridical sphere overlap in the common sense and, hence, in their visual representations.

Medium rated icons : these icons are illustrated in *see* Fig. 6.17. **Consent** was expressly symbolized as a choice between accept and refuse (as opposed to more common passive acceptance). Although half of the respondents noticed that emphasis was put on the possibility, i.e. right of choice, between accept/refuse or agree/disagree, which corresponded to the designers' intention, the lower rates are due to the fact that the tick and the cross can be ambiguously associated also to the dichotomy of right/wrong, yes/no, true/false. Although it is reasonable to assume that the provision of more contextual cues would shrink the number of possible interpretations, this must be proved. Similar critics received the icon for **the right to withdraw consent**, based on the same symbols, combined with an arrow to signify the transition from given consent to withdrawn consent, which was however correctly understood by almost all the respondents.



(a) Consent

(b) Right to withdraw consent

(c) Legal obligation

(d) Vital interest

(e) Right to data portability

(f) Right to rectification

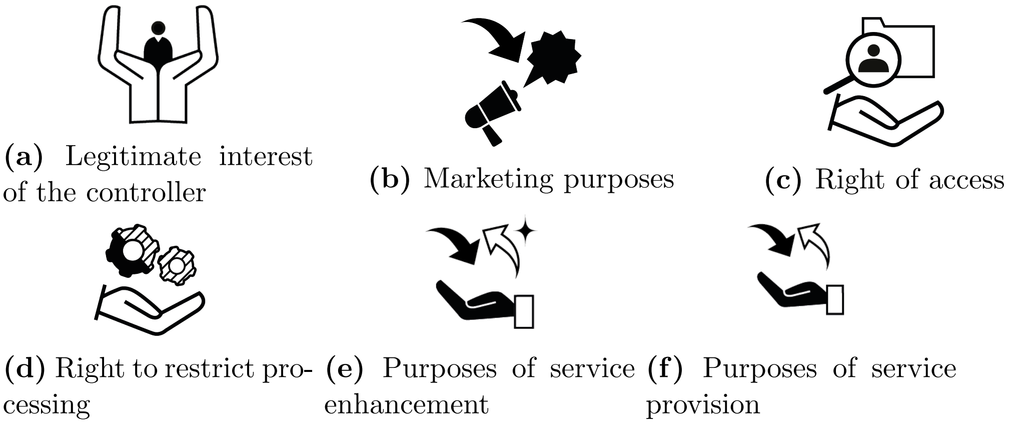
Figure 6.17: The medium rated icons in Task 2

Mixed opinions have been gathered on the icon for **legal obligation**: whereas a few endorsed the stamp as symbol for official, thus per extension legal, some others expressed their doubts since the stamp can be also linked to the administrative domain, such as a certificate, and a few also suggested to use the hammer to be more legally specific. The variety of opinions gathered and the good arguments provided suggest that in order to be more easily and unambiguously understood, more appropriate alternatives for this icon should be created and tested. The graphical symbol for **vital interest** also received comments of mixed nature: although the EFC conveys the idea of vital importance, thus concerning life and death, proposed enhancements would possibly specify the icon in context. Indeed, “it can look like an audio file” (P16) and “it can be confused with the device’ life” (P9). In addition, it could be too strictly linked to the health domain because it’s “the health metaphor for anyone” (P3).

The folder in shape of a suitcase, a metaphor symbolizing the **right to data portability**, was positively embraced by three quarters of the subjects. However, a few respondents specified that the transfer of data from an entity to another should have been better specified, e.g. with arrows.

Although the marks given to the **right to rectification** are not high, which contradict the researchers’ expectations, the explanations from all the participants mention the pencil as a clear symbol for modification (borrowed by the many applications that use the same symbol for the edit function). However, one fourth of the respondents expressed the need to specify the object of modification, e.g. by adding a data folder (P18), whereas at least in two cases the low grade was given because it was the concept of “rectification” that was not grasped, but, once explained, the symbol was considered appropriate.

Worst rated icons : At the lower end of the spectrum, the icons that were rated more poorly (*see* Fig. 6.18), for instance the **legitimate interest of the controller**. Although one fourth of the respondents noticed the controller’s specific clothing referring to the “higher social status of the person” (P7), “an authority” (P14), “elegantly dressed, so maybe in a position of power" (P18), more than one third could not distinguish the controller from the user - a legibility problem already highlighted in the first part of the test. These low rates might also be due to the two hands to signify the interest (but *see* discussion below). The icon for **marketing** was also poorly rated, mainly because it was deemed too similar to any other kind of communication, thus too general to be exclusively attributed to the advertisement sector. In a couple of cases it was wrongly interpreted as dissemination of personal data.



(a) Legitimate interest of the controller

(b) Marketing purposes

(c) Right of access

(d) Right to restrict processing

(e) Purposes of service enhancement

(f) Purposes of service provision

Figure 6.18: The worst rated icons in Task 2

The **right of access** also got mixed marks. On the one hand, half of the participants correctly interpreted the magnifying glass as metaphor for looking into the folder, thus accessing the files. Criticism was raised, however, on the fact that it is unclear that it is the data subject's data (and not someone else's data) that s owned by another subject and not by the data subject (*see* the discussion below on this point).

The **right to restrict processing** ranked among the worst icons for a series of reasons: firstly, half of the participants underlined how the gears' symbol is usually employed to signify the device settings, so it does not unequivocally recall the processing. Secondly, the idea of restriction/limitation symbolized by the difference in gears' color was not easily grasped, either because the difference was too subtle or because it could not be traced back to its meaning. Nevertheless, half of the participants could understand the designers' intentions, even those that rated the icon poorly.

The icons for **purposes of service enhancement** and **purposes of service provision** that ranked second to last and last, respectively, can be discussed together since they are almost identical and therefore present similar problems. Although the idea of an exchange represented by the two arrows in opposite direction was approved by more than half of the subjects, the icon was described as "incomprehensible", "not intuitive", "unclear", "vague" and a few respondents clarified that they would need an explicit explanation. However, the star/plus element was positively rated and easily interpreted by three quarters of the subjects as symbol for enhancement.

6.5.5.3 Task 3: Alignments with Designers' Intentions

Figure 6.19 displays the results for each visual element appearing in DaPIS that has self-contained meaning. The best results (e.g. a group of users meaning public, a pencil that stands for rectification, a shield signifying security, etc.) seem to reproduce the icons that scored better in the previous task, whereas the elements towards the end of the graph (e.g. the column, the controller's black hand with white sleeves, the rights' hand, the gears for processing) were assigned more frequently a wrong association with the designer's intended meaning, thus indicating misinterpretations.

6.5.5.4 Task 4: Best Alternative

Among the two icons produced for the right to object to processing, three quarters of the respondents preferred the icon with sharply separated gears because it could more easily suggest a complete break that can not be fixed.

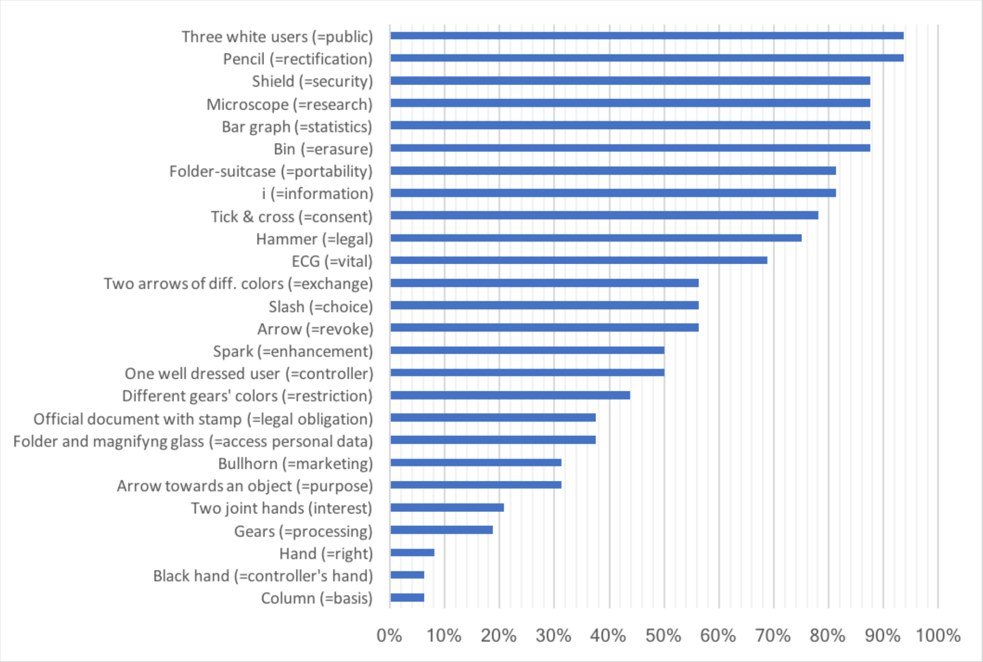


Figure 6.19: Percentage values of correct explanations for designers' iconographical choices

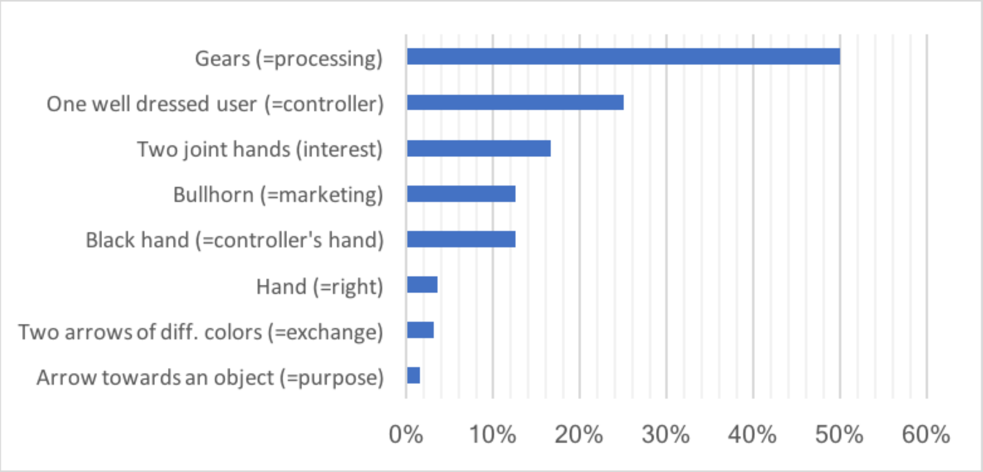


Figure 6.20: Percentage values of wrong tentative explanations provided for designers' iconographical choices. For the elements not shown on the graph, no wrong explanation was provided

6.5.6 Discussion of Results

Legibility. Concerning legibility, the few icons reported below need to be enhanced, in order of severity:

1. right to lodge a complaint to a supervisory authority: there are too many elements, some of which are too small to be effortlessly recognized;
2. legal basis: either make the column more recognizable (it is hard to tell whether the metaphor for basis was understood due to its low legibility) or take it away;
3. right to data portability: the folder needs to resemble more closely to a bag or a suitcase in order for the metaphor to be grasped;
4. controller: as some comments in the recognition task also confirm, its difference from the data subject's silhouette needs to be more prominent, e.g. by widening its white parts or adding additional distinctive marks;
5. vital interest: make the ECG more distinctive or provide some contest around it.

Furthermore, some graphical symbols were too emblematic for being described in terms of their components. For instance, the stars in circle were identified straightforwardly as the EU stars and described in those terms. Also the tick and the cross received a similar treatment.

Concreteness and Familiarity. Not surprisingly, the icons that scored best represent concrete objects, familiar concepts or are based on familiar representations (e.g. information, erasure). This reality is also reflected by the users' explanations displayed above. Conversely, the concepts behind the icons that scored worst, e.g. for the provision or enhancement of the service, are vague, general, and abstract. During the design phase, close scrutiny and long discussions originated around possible ways of representing these notions. Although for such concepts a semantically transparent solution might never be reached, alternatives should be explored and compared to this version of DaPIS to determine whether they are more readily graspable. Indeed, in the last test that will be illustrated in the following, some icon alternatives for these notions were proposed and evaluated. The analysis of the frequencies also indicates that, whereas for better ranked icons judgments were more compact, marks are distributed along the five possible marks as the icons become less familiar or less concrete. Individual characteristics might be the cause.

Interpretation of Iconographical Choices. A general tendency that can be noted is that, in those cases where the association between icon and referent was deemed more appropriate, the users' explanation more frequently coincided with the reason behind the design choice. For instance, participant P2 explains her high mark for the right to be informed with the fact that the "i" is an unquestionable symbol for information that everybody knows and this is why the designers chose it. For what concerns symbols that were rated poorly, in some cases the designer's intention was nevertheless understood and explained. For instance, another participant (P3) expresses her doubts on both icons for purposes of service provision and of service enhancement, by saying that they are not intuitive and that she would not grasp the underlying meaning from the image alone. Nevertheless, she is able to provide correct and accurate explanations for the designers' choices: the arrows signify an exchange, whereas the star stands for enhancement. Also the explanation provided for the right to data portability is exemplary in this respect: "I would not click on this icon to receive or transfer my data. Intuitively, I would have accepted a symbol of entrance/exit" she provides as reason for her mark expressing incertitude. However, when asked about the supposed line of reasoning behind the folder, she effortlessly identifies the folder with the handle as metaphor for transportability. An additional emblematic case is presented by the icon for the right to restriction of processing: notwithstanding the poor evaluation given to the difference in gears' color, meaning limitation, almost half of the participants could explicitly and correctly associate it with its intended meaning. This seems to indicate that, even if some visual choices are not readily grasped, some consideration can guide the interpretation process and align the mental models. This is also shown explicitly by some users' comments, e.g. when P19 notices the recurrence of the joined hands to signify the interest: "it is not easy to link the hands to the concept of interest. [But] once established that the interest is represented by the hands in this position, then it is easy to identify the controller in the picture". In other cases, however, the participants were confused about the reasons behind certain iconographical choices and could not follow the line of reasoning: for instance, the reason why the icon with gears has been selected to signify processing, whereas it is the usual icon for settings, or the reason why the joint hands signify interest.

The Hands' Meaning. The use and interpretation of the hands symbol must also be commented. The hand with the palms facing up (i.e. the "holding hand") has the metaphorical extension of "being in control" or "have the power over" to indicate the possibility granted by a right to its holder. Even in the legibility phase multiple interpretations were offered for the symbol: a offering hand (P2), a welcoming hand (P3), a helping hand (P14), a requesting hand (P17), a hand offering a possibility (P9 e P14), a protecting hand (P17). During the second task, less interpretations were provided since the participants focused more on the distinctive element placed above the hand, than on the hand itself. Only four individuals explicitly referred to the hand's intended meaning and all of them after a few exposures, thus after having noticed that it is the recurring element to signify a right. Others seemed to interpret the hand by drawing inferences from the other elements of the icon or from the definition provided, e.g. for the right of access:

P2 provided the explanation “me, data that I own”, P10 referred to “grab to open”, for P15 it signifies “data in our hands”, whereas P17 said that it represents “someone external that holds that about you”. Although the hand symbol is rather arbitrary, thus variations of interpretation are inevitable, the metaphor seems not to have been smoothly decoded. Alternatives could therefore be researched and compared to this icon to find out if they perform better. The joined hands to signify interest were also oppositely interpreted: either they were recognized as metaphors for protection or care for the elements contained in-between, or as synonym of power exerted over the elements. For this reason, alternatives to this symbol have been explored in the following design workshop and user testing, alongside with some minor modifications suggested during this study (*see* also Section 6.6).

Usability versus Precision. The comments on the icon for the right of access, i.e. that it is unclear that it is the data controller that has the data that the data subject wants to access, seem to echo the reasons that brought to the literal representation of this concept in the first icon subset (*see* Sect. 6.4.6). The same conclusion, however, holds: some details need to be sacrificed and left to textual provisions for the sake of icons’ usability.

Re-use or Invention of New Icons? Many comments, even from the legibility testing phase, highlighted that the gears in isolation are immediately linked to settings rather than processing. This metaphor has been consistently used across the icon set, although it must be reckoned that the original version saw a composition of arrow and gears to signify a transformation of personal data into something else: some kind of processing, indeed (*see* also Fig. 6.7). It must be therefore researched if some element must be added to specify the gears’ meaning or, rather, if a completely different metaphor must be found. This example shows that it is challenging to strike a balance between the reuse of known and deployed symbols to ease understanding of new concepts and the use of new symbols. The first case entails the risk of extending (or maybe overstretching) the use of a symbol (e.g. gears) that is already strictly linked to a primary meaning (e.g. settings) to a different meaning (e.g. processing in DaPIS), while resorting to new symbols can prevent this over-extension. However, the acquisition of a completely new visual alphabet might be characterized by a steep learning curve and might happen with great effort.

Familiarity with Concepts. Finally, there is a specific case that can shed light about how previous knowledge of concepts can positively influence icon understanding. During the legibility task, one of the study participants freely provided correct interpretations of a number of icons: not only on the more familiar ones (i.e. right to be informed, right to erasure of data, data transfer to third countries), but also on less immediate icons, such as right to data portability, data controller, encrypted data, pseudonymized data. Although the participant described herself as having intermediate legal knowledge, her answers clearly indicate an accurate knowledge of the topics.

6.5.7 Limitations of the Study

The participants in the two user studies were really diverse, as well as the types of tasks that they carried out. For these reasons, the results are not comparable. In the second study, although a greater variance was expected, almost the entire totality of the subjects were Italian individuals in their twenties with a high educational background. However, involve very wide, differentiated, international audiences that are representative of European population is admittedly out of our reach (*see* also Section 6.8).

From this also follows that, as in other lab researches, the tasks of this user study presume a serious consideration over icons that might not closely mirror the presumably quick sense-making process carried out in real-world conditions. However, such a limitation can be overcome only if organizations start to employ the icons in different contexts (online versus offline, paper versus digital, in combination with text versus as stand-alone elements, etc.) and "in the wild", i.e. on online platforms, social networks, etc.

Moreover, since the interpretation keys were provided, it was impossible to explore the informational value of the icons and it is plausible to assume that some icons' received a better score that they would have received in a classical matching task. Although on the one hand this strategy partially simulated the research activity of a user that wants to find a specific piece of information that she already knows (i.e. from concept to icon), on the other hand it is the opposite of facing unknown symbols for the first time and use them to infer their meaning (i.e. from icon to concept). Both directions must be researched.

Finally, it is well-known in literature that asking for marks in-person sparks higher grades, but this format was chosen because it would have been otherwise impossible to encourage explanations about the mark. To counterbalance this effect, it was made clear that negative marks were very useful for the research.

6.6 Evaluation of the Third Icon Design

6.6.1 Introductory Considerations

After the last redesign of DaPIS (*see* Section 6.2.2.3), it was deemed necessary to run a further evaluation of these icons, on the model of the last one. Moreover, some of the icons of the first icon set, for example those related to the processing operations, had not been previously evaluated.

It was decided to keep a similar pool of users, i.e. highly educated young people (20-35). For this reason, motivated individuals owning at least a master degree were recruited across some universities, mostly at the University of Luxembourg, which also guaranteed a more international audience than the two previous tests. Another constraint was a high level of English proficiency, to ensure easy comprehension of the questions and of the legal definitions, and to ensure sufficient linguistic means to provide detailed and elaborated answers. The study was carried out in an online environment to be more easily distributed to the participants and also to experiment if, given the lessons learned from the previous studies, it is possible to conduct such a test online and at distance, within the view of future large-scale distribution for a final evaluation (*see* next Chapter). The organization of in-person studies, in fact, and the collection of results are usually extremely time-consuming and unfeasible on large scales.

6.6.2 Participants

10 participants took part in this online study, all having at least a master degree and advanced English level. All the participants described themselves as having intermediate or advanced digital competencies, whereas their legal competences are placed on two opposites, i.e. either advanced or basic. Their origins are Italian, Armenian, Iranian, Canadian and Greek. Non-EU residents have lived in Europe at least since one year. Their age ranges from 28 till 33, with an average of 31 years old.

6.6.3 Tasks

For the reasons anticipated earlier, this last user study was carried out online, on the website that documents the research[10], in July 2018. This study replicated the same kind of tasks of the second user study (*see* Section 6.5): subjective rating on the correspondence between icon and concept, and alignment with designers' intentions for those icons that had not received previous evaluation or that were completely redesigned based on the results of the last user study. Moreover, it was asked to choose among two or more alternative icons for the same concept in 5 cases, i.e. the more problematic cases in terms of simplification versus completeness of representation (as in

[10] http://gdprbydesign.cirsfid.unibo.it/

the icons for contract and legal obligation) or in terms of comprehensibility (as in the icons for provision of a service, sharing with third parties, and the hands symbol for interest as in public interest).

Since the study was not undertaken in the presence of a researcher, it was fundamental to provide very clear explanations for the tasks, especially about how to provide meaningful answers. Illustrative answers were thus provided at the beginning of the test and detailed explanations about the desired answers were given, e.g. "Please try to provide a precise answer that will help us understand what you think and why you think it. Avoid general answers like 'because it's clear/not clear' ". More importantly, as previous experience shows, the icons are meaningful only if considered in context: firstly, the icons and their elements have to be understood as part of a set because only in this manner some elements become understandable (e.g. the recurring hand symbol to signify the data subjects' rights); secondly, the icons' function as information-markers in privacy policies must be made clear. This is why the icons were displayed in groups according to their conceptual category and the relevant section of privacy policy (see H) was displayed right above the icons pertaining to that section to provide enough contexts to the respondents. The choice to display the mock-up of a visualized privacy policy was also motivated by the need of reproducing similar conditions to the second user test (*see* Appendix A). Early feedback about the questionnaire design and the questions' wording from two colleagues helped to set up a smooth experience for respondents that had no familiarity with the icons and the typology of questions.

Three tasks were designed:

Task 1: Subjective rating on the icons' correspondence with its underlying meaning: a subjective rating on a Likert scale that ranged from 1 to 5 about the ability of a certain icon to represent the correspondent concept, expressed through a label and a simplified definition, was chosen. Since the researcher was not present, an explanation for the marks was provided: 1) Strongly disagree (you find it impossible to associate the icon with its meaning); 2) Agree (with some changes or efforts of interpretation, the icon could work); 3) Neither agree nor disagree (you do not have sufficient elements to express your opinion); 4) Agree (the icon can work, but it needs minor improvements); 5) Strongly agree (you could not think of a better icon for the concept). Explicit explanations for the mark were asked and recorded in written form in a dedicated space below.

Task 2: Alignment between users' mental models and designers' mental models: the participants were asked to attempt to provide an explanation for the visual choices made by the designers. This task was meant to find out whether users could understand the reasons behind the choices and thus align their line of reasoning with the designers', despite their opinion on the appropriateness of a certain icon for a certain concept, evaluated in the previous task. In order to avoid influence by different wordings, the same identical simplified definitions that were distributed to the designers to spark the design process, were also provided to the test participants.

Task 3: Best Alternative: five alternative choices among two or more icons were asked.

6.6.4 Analysis

The data collected in written form by the participants was gathered and analyzed by one of the interviewer in search for common patterns. Given the nature of the tasks and the study goals, qualitative analysis was the main source of data. 14 icons were evaluated: whereas some underwent assessment for the first time, for some others it was the best alternative among multiple options the focus of the assessment.

For the fitness of correspondence between icon and concept for the results of data processing operations (i.e. “anonymized data”, “pseudonymized data”, “encrypted data”, “profiling”, “automated decision-making”, “storage of data inside of the EU") and for the rights of the data subject (i.e. "rights of the data subject", "right to object to processing", "right to lodge a complaint to a supervisory authority") (task 1):

Average : mathematical mean across the users' self-reported marks was computed for each icon;

Furthermore, the explanations provided by the participants to motivate their choice were analyzed to find common lines of reasoning, but also to find out explicitly about the words that were employed to describe why some associations were easier/harder than others.

For the alignment between users' and designers' mental models of all the icons, i.e. the ones listed above in task 1 and below in task 3 (task 2), hit rate and error rate were computed as in the second user study (see Section 6.5).

For the best alternative between the alternative icons for the concepts of: "contract", "legal obligation", "public interest", "purpose of provision of the service", "data sharing with third parties" (task 3) (*see* Figures in Section 6.6.5.3):

Higher number of preferences : the best alternative was chosen simply by counting which icons got the majority of votes. Reasons for the choice were also recorded.

6.6.5 Results

6.6.5.1 Task 1: Rating on Fitness of Correspondence

The results of the assessment on the icons' capacity of representing the underlying concept are reported in Fig. 6.21. The results reported in the following paragraphs are organized in three groups, according to the average value obtained by each icon:

1. best rated icons: average value ranging between 5 (=completely agree with the fitness of correspondence between icon and concept) and 4 (=agree);

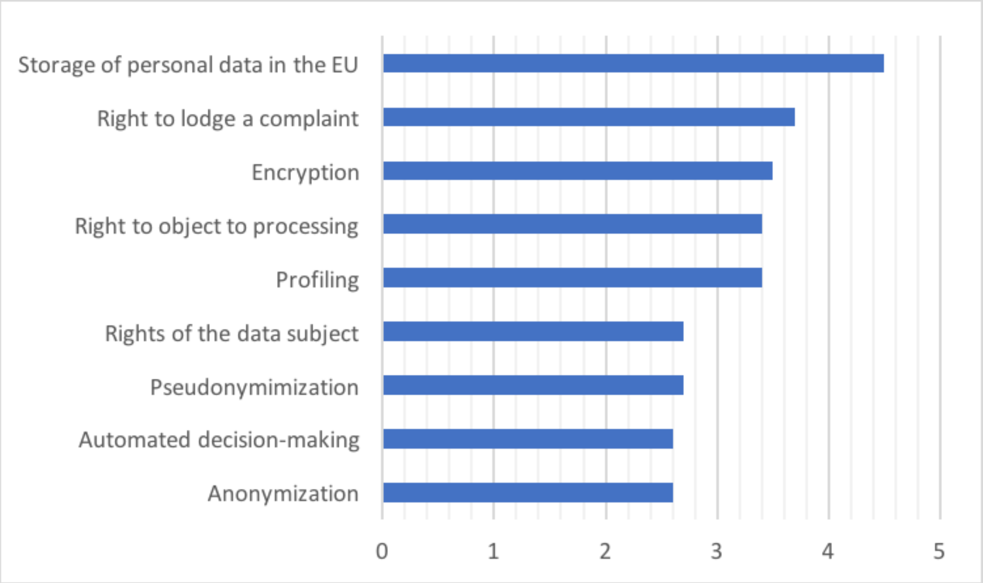


Figure 6.21: Means of the self-reported values for ease of understanding, where the closest to 5, the better the results

2. medium rated icons: average value ranging between 4 (=agree) and 3 (=uncertain);

3. worst rated icons: average value ranging between 3 (=uncertain) and 2 (=disagree).

Best rated icon: the only icon that received very positive ratings is the symbol representing the **storage of data in the European Union** (*see* Fig. 6.22), which is the counterpart of the transfer of data outside the EU. Every participant was able to recognize the stars in circle as the "EU flag" or "the symbol of the EU" or similar, while the personal folder placed in the middle was understood as data "physically stored inside the European Union" (P4).

Medium rated icons: there are four icons that received marks between "I agree" and "Neither agree nor disagree" (Fig. 6.23). The icon for **right to lodge a complaint to a supervisory authority** was quite positively evaluated because a person sitting on an armchair and behind a

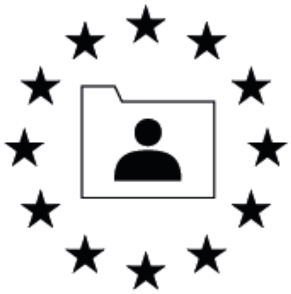
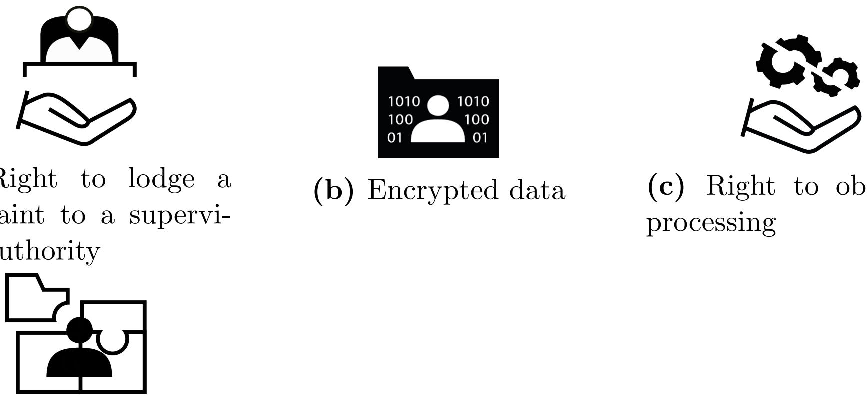


Figure 6.22: The best rated icon in Task 1 is the storage of data inside the EU

table was seen as "person with power or authority" (P2) or "someone who is responsible and can make authoritative decisions" (P4). **Encrypted data** ranked next, because the combination of zeros and ones was interpreted as "content that is not readable by everybody" (P4), "binary language that is symbolic [...] and expresses a code" (P5). Those participants that gave lower scores mainly appointed it to the fact that they would have expected a padlock (*see* Section 6.6.6). As for what concerns the **right to object to processing**, the gears were positively welcomed as expressing processing (although a few participants pointed out that they are more readily associated to settings), while the fact that they are broken was also easily understood as interruption, although a few people also suggested alternative symbols (*see* discussion). Similarly scores received the icon for profiling, with participants appreciating the idea of the puzzle pieces relating to "different aspect of personal information" (P1) through which "it is possible to reconstruct the individual identity and preferences" (P2) or "reconstruct the behaviour" (P10) in a process of "profile-creation" (P3). However, four participants pointed out that the icon could be interpreted as "decomposition" (P9) of the data folder, as if the pieces were "separated" (P5), instead of composed together.

Worst rated icons: four icons received grades between "Neither agree nor disagree" or "Disagree" (*see* Fig. 6.24). Although the diamond in the



(a) Right to lodge a complaint to a supervisory authority

(b) Encrypted data

(c) Right to object to processing

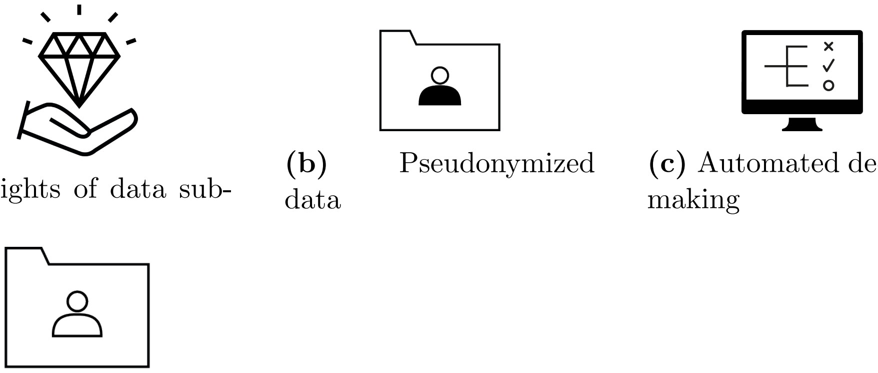
(d) Profiling

Figure 6.23: The medium rated icons in Task 1

rights of the data subject was correctly interpreted as “my rights are important” (P10), “something precious” (P8), “invaluable right” (P4), “something valuable and important” (P2) or ”relevant” referring to the rights, not everybody understood the metaphor, while the hand was interpreted with difficulties. The icon for **pseudonymized data** and **anonymized data** also received similar low marks, mostly because the symbolic differences of colors and the icon representing the personal data folder was grasped, but deemed difficult to readily associate with the intended meaning and to distinguish among similar icons. Lastly, **automated decision-making** was deemed very difficult or impossible to associate to its underlying concept, with explanations like “it takes a good effort” (P7), “I cannot *see* neither the ‘decision-making’, nor the ‘automated’ concept” (P9), “I do not *see* the connection” (P10).

6.6.5.2 Task 2: Alignments with Designers’ Intentions

Figure 6.25 displays the results for each visual element with self-contained meaning appearing in the batch of icons analyzed in this last phase. The EU flag stars, the changing colors of the user’s silhouette and the puzzle pieces were more correctly associated to their intended meaning, whilst the designer’s intention behind the hand, the authority, and the computer was



(a) Rights of data subjects

(b) Pseudonymized data

(c) Automated decision-making

(d) Anonymized data

Figure 6.24: The worst rated icons in Task 1

not readily grasped. The results also show some errors of interpretation that will be discussed in Section 6.6.6.

6.6.5.3 Task 3: Best Alternatives

Following the dismal results from the second user study, two alternative icons were designed for the concept of **contract** (*see* Fig. 6.26): one that more precisely represents the legal relation between the user and the controller (Fig. 6.26a) and another that, for the sake of usability, only represents the written agreement (Fig. 6.26b) . Seven out of ten respondents preferred the representation showing the legal bound between two entities because the document alone "could be anything" (P7) and "not necessarily a contract" (P4), but "could be a simple letter" (P5).

Since the second user evaluation showed that the icon for the **legal obligation** (Fig. 6.27a) was deemed too similar to a certification and not enough specialized in the legal sense, two alternative icons were tested together with the original one (*see* Fig. 6.27). The original idea for the icon, that had been simplified for usability reasons, was recovered: a pointing hand was thus added to the icon in two different versions, a simpler icon without the stamp (Fig. 6.27c) and a more elaborated one with it (Fig. 6.27b). The preferences of the respondents are distributed almost evenly among the three alterna-

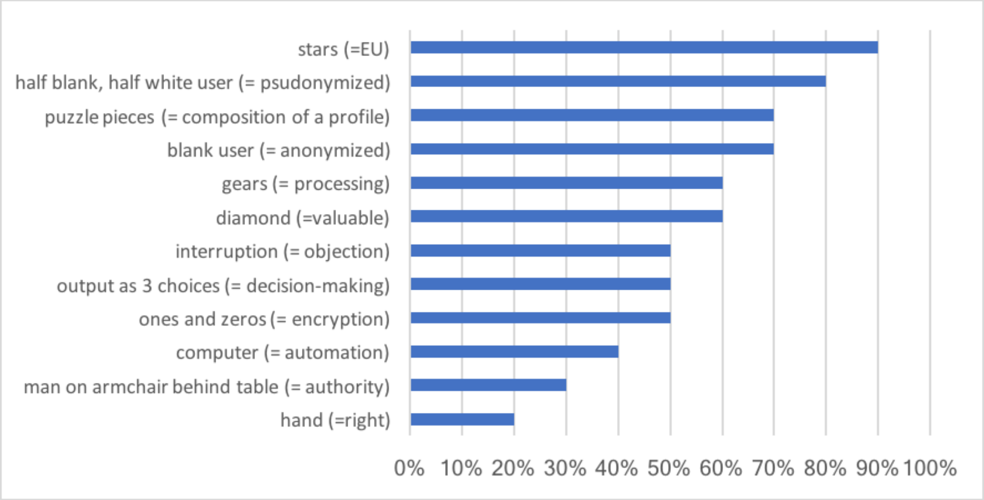
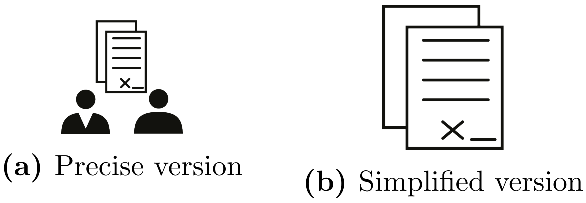


Figure 6.25: Percentage values of correct explanations for designers' iconographical choices

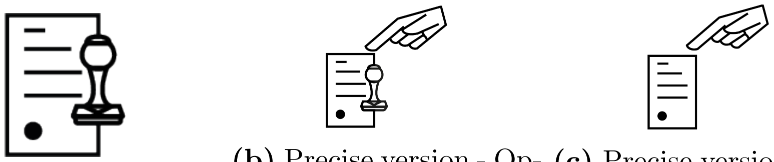


(a) Precise version

(b) Simplified version

Figure 6.26: The two alternative icons representing a contract

tives: three respondents preferred Fig. 6.27a because "a sealed document shows an obligation" (P6) and "the hand in the other icons was not meaningful" (P9); four people elected Fig. 6.27c "because of the hand (authority) and also a hand pointing at the rules making it look stricter" (P8) and "the rubber stamp gives me the idea of a 'legal' process" (P10); finally, three participants chose Fig. 6.27b, citing similar reasons for the hand (e.g. "an external intervention (law in this case) that obliges me to do something" (P2) and "the hand makes me think of something mandatory" (P5)), but the stamp recalls a registered contract (P2), is unnecessary (P4) or is unrecognizable (P5). The combination of the answers suggest that the pointing hand indicating the obligation can be meaningful, while the details of the sealed document can be simplified, as long as the symbol is not mistaken with another official document (e.g. a contract, a certification, etc.).



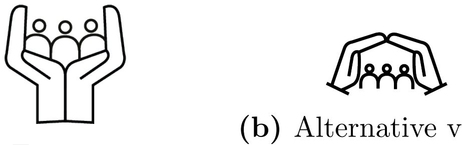
(a) Simplified version

(b) Precise version - Option A

(c) Precise version - Option B

Figure 6.27: The three alternative icons representing a legal obligation

The third symbol that was investigated in this phase is the hand related to **interest**, which is the basic element that, combined with others, generates the icons representing "public interest", "interest of the controller", and "vital interest". In order to focus participant's attention on the hand element, the evaluation was carried out on the example of "public interest", which had shown less interpretative issues compared to the other two in the preceding user test. The original icon (Fig. 6.28a) was thus compared to an alternative version with the hands placed in a different manner (Fig. 6.28b). The original version gathered seven preferences out of ten, with the symbol of hands interpreted as "supporting a group of people" (P2), while the second icon was interpreted univocally as "protection" by all participants, a concept not corresponding to the "promotion" of the public interest (P5). This clearly means that, although the meaning of the "supporting hands" might not be totally transparent, the "protecting hands" is certainly not a good alternative.



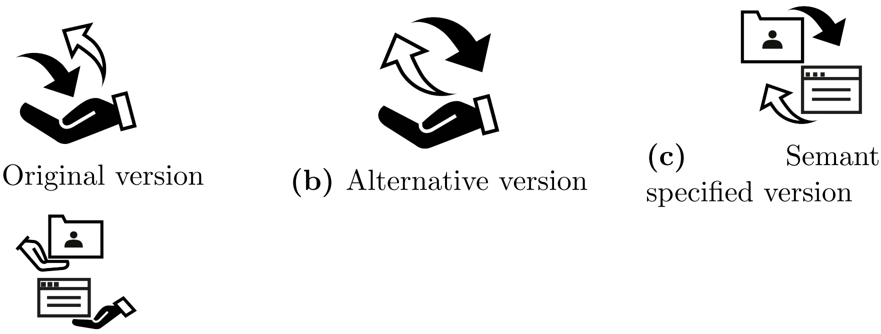
(a) First version

(b) Alternative version

Figure 6.28: The two alternative icons representing the public interest

As the discussion of results in the last section highlighted, one of the most controversial graphical symbols was the one attempting to represent the **purposes of provision of the service** (Fig. 6.29a) and the very similar symbol for enhancement of the service. By taking into account the comments received in the preceding user test, three alternatives were designed: one icon closely resembling the original, but with arrows suggesting a circular movement (Fig. 6.29b); one icon that is more semantically specified, with arrows signifying the exchange between personal data and a service, exemplified by a webpage (Fig. 6.29c); and lastly, a similar icon that, instead of the arrows, re-uses the contrast of colors between hands adopted by other symbols pf the set to signify the data subject providing data on the one side and the controller providing the service on the other side (Fig. 6.29d). The preferences are evenly distributed between the two icons that only contain arrows and the two icons that more specifically represent the exchange of data with a service. The two participants selecting the original icon motivated it with the fact that it is more generally representative, while the three people preferring the other circular disposition of the arrows either did not provide any meaningful explanation, or appreciated the order of the two arrows because "we often give before we receive [...] which is often the case when a service is provided" (P5). The icon in Fig. 6.29c received four preferences because of its precise and concrete nature, but was also deemed quite complex, while the fourth choice only received one mark in opposition to the others in which the arrows "are too related with recycling images". In conclusion, the debate around the best manner to represent this concept is still open.

Lastly, three alternative icons for the concept of **data sharing with third parties** were evaluated: one icon based on the standard symbol of sharing used nowadays on many applications, combined with three parties (Fig. 6.30a); a second icon representing a simplified globe (Fig. 6.30b); and a third icon representing three interconnected parties (Fig. 6.30c). The icon employing the popular symbol of sharing received seven out of ten preferences, precisely because the symbol is "well-known" (p8), "familiar" (P1, P3) and "already in our common understanding" (P2). The alternative representing the globe was chosen in one case because it highlights "the possibility (danger) of data to travel to other (far-away) parties [that] may include a sig-



(a) Original version

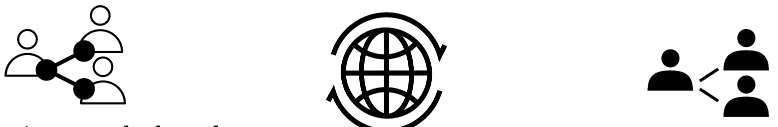
(b) Alternative version

(c) Semantically-specified version

(d) Semantically-specified version

Figure 6.29: The four alternative icons representing the purposes of provision of the service

nificant loss of control by the user" (P9), while the third icon best represents a "network of people" (P5).



(a) Sharing symbol and three parties

(b) Global reach symbol

(c) Three parties

Figure 6.30: The three alternative icons representing the data sharing with third parties

6.6.6 Discussion of Results

Future Redesign It was expected that one of the best rated icons would have been the one representing storage of data inside of the EU, because of the familiarity with the stars of the European flag. The marks for the icon representing the right to lodge a complaint are satisfactory because the design of the supervisory authority witnessed many iterations. However, some participants pointed out that the fact that there is the possibility to make a complaint is not understandable. For sure, an alternative showing a padlock for encrypted data will need to be explored because, as some people stressed in their answers, it is a more recognizable symbol for the ordinary user than the metaphor of zeros and ones, that is more correct but probably less transparent for people without technical experience.

The icon for the right to object to processing received similar comments to the icon for the right to restriction of processing: the gears have been more easily associated with processing than in the previous experiment, one reason probably being that the participants were more international. However, some comments pointed out that they can be confused with settings. Moreover, a few suggestions made reference to the fact that the objection could be symbolized by the symbol for a stop/alt, because the broken gears might rather suggest that there are difficulties with the processing. It was also expected that the representation for the rights of the data subject could have been easily misunderstood, given the abstractness of the concept and the metaphor behind the diamond. This symbol was easily linked to something "valuable and important for the user" (P2), but it was also pointed out that it "makes the assumption that people care about their rights" (P7), which is not always the case. Nevertheless, this design choice was deliberately metaphorical and positive to attempt to convey the importance and value that rights can assume for data subjects.

Contrary to the researchers' expectations, the icon for automated decision-making was not well rated. The comments reveal that, although the decision-making process was more easily grasped, the computer symbolizing the automation was not understood. Some participants suggested to explore graphical symbols concerning robots, because more easily associated to automation, or to stress the absence of human intervention.

Directionality of Interpretation Some comments on the graphical symbol representing the concept of profiling also require attention: a few respondents were undecided on whether the puzzle pieces are gathered together to compose the personal data folder or, vice versa, if they are separated from each other in a process of decomposition. Indeed, the image can be read in both senses and icons from the first icon set had received similar comments. However, as it will be discussed in the next chapter, it is hard to show the direction of a process in a static image, whereas movement can be easily conveyed by a gif or any other sort of animated visual.

Black and White Colors The icons for anonymized and pseudonymized data scored badly, but the answers of the respondents reveal that, once that the difference among the colors is noticed, than the metaphor behind the colors is understood: a black user silhouette to identify personal data, a blank silhouette for anonymized data, and a half blank and half white silhouette for pseudonymized data. However, two considerations are necessary: firstly, the meaning of the colors can be grasped only if each icon is shown in combination with at least another icon, as also some respondents hold; secondly, from this derives that such difference might be too subtle to be readily understandable. Many participants also underlined that an additional difficulty had to be ascribed to the complexity of the concept itself.

Usability versus Precision On a general note, the tension between simplicity of design, relevant for usability reasons, and preciseness of representation, important for legal reasons, re-emerged prominently also in this user study. Icon alternatives for the concept of contract, legal obligation and provision of the service were redesigned exactly because in previous experiences the need for more precise representations had emerged. As it will also be discussed in the next chapter, however, this opposition is not easily solvable and can also depend from individual preferences: comments from this last user study show that whereas for some individuals the precision of representation of a concept is of utmost importance, for others simplicity over the complexity of design has to be favored. For example, compare the comments of two respondents that motivate their choice of a contract icon over the other: whereas P4 motivates her choice as "I prefer the icon with the two silhouettes in it because it conveys the fact that this document (contract) is agreed upon between the two parties. The document without any silhouettes in it could illustrate a set of rules and not necessarily a contract", P10 writes: "I think the simpler the better. The two users are redundant". Hence, more research on this point is needed since this question has re-emerged persistently during every workshop and user study.

One more point that deserves discussion is raised by the representation of the concept of data sharing with third parties. Whereas the two icons displaying three users more closely and literally resembled the underlying concept, and for this reason were preferred by the majority of respondents, the icon displaying a globe focused on the fact that data can be scattered, without exact knowledge of its recipients. In this light, this icon can also, to a certain extent, convey the risks inherited by the data sharing.

Semantic Specifications Another recurrent comment concerned the absence of the folder representing personal data as building blocks in other icons, as in the icon representing the sharing with third parties. Some users would have expected to *see* the application domain, i.e. personal data, in the icons, otherwise the visual elements seemed to them not enough determined, i.e. what is shared with third parties? A similar discussion was opened up in the last user study, with some participants calling for the presence of the folder representing personal data in, for example, the right to rectification because according to them it was otherwise impossible to determine what this right concerned, i.e. the scope of rectification.

Although these comments calling for a semantic specialization in the domain of personal data are reasonable, two considerations are re-

> quired: firstly, it is the privacy-related context (e.g. the privacy policy) that provides this semantic specialization to the icons and makes it redundant to specify it with visual elements that would make the icon unnecessary crowded. Secondly, for the very functional nature of icons, it is impossible to specify meaning to that extent: as it will be advocated in the next chapter, other elements (like pictograms or comics) are more suitable for an exact and detailed representation of meaning.

6.6.7 Limitations of the Study

Given that it reproduces the same method of the second user study (*see* Section 6.5), also this third user study presents similar limitations concerning the fact that it does not reproduce real-world conditions of icon interpretation, since the participants had the time to scrutinize the icons with attention and the corresponding concept was provided. Nevertheless, this study, like the previous one, only attempts to gather information on people's sense-making process and to offer subjective evaluations about the fitness of correspondence between a symbol and its concepts.

A major limitation constituted by the representativeness of the pool of participants, that were only ten, young, and well-educated, which does not correspond to the majority of Europeans. Although their demographics was expressly selected to be similar to the participants from the previous study, it can be the case that their ratings on the icons excluded in that study would have been different. The fact that the evaluation was not done in the presence of a researcher might as well have influenced the results, because the participants might have felt more free to express negative judgments.

Finally, since this was the third iteration of the icon evaluation, only a selection of icons was shown to the study participants: those that had not received any previous evaluations and those that had been vetted after the last user study. It was attempted to counterweight this limitation by showing the context where the icons would appear, e.g. the privacy policy's section containing the icons of the same conceptual category. However, in the next evaluation phases, it will be fundamental to consider all the icons at the same time, because the interpretation of one icon can be built on and supported by the previous interpretation of a similar icon, as the example of the category of data subjects' rights shows.

In conclusion, the results exposed in these pages only intend to provide a preliminary indication of the more promising icons. Nevertheless, more experimentation is needed and further considerations on this point will be provided in the next chapter.

6.7 The Final DaPIS

According to the results discussed in the previous pages, in E are displayed the icons composing DaPIS. It can be safely assumed that most of these elements can be adopted, although there still are few icons that need further research and these are specified by an asterisk. In the next Chapter the results and limitations of the research described in this chapter will be discussed on a more general level in order to recommend future directions of research. The final DaPIS can be downloaded from the research website[11] and is licensed under a Creative Commons Attributions-ShareAlike 4.0 International License[12].

6.8 Conclusive Remarks

6.8.1 Necessary Efforts towards Standardization and Education

As other researchers before us [125, 91], we also end with a note on the necessity of standardization, since there will always be a margin for individual, i.e. free, interpretation of the icons. Whereas researchers from

[11] http://gdprbydesign.cirsfid.unibo.it/dapis-2/

[12] https://creativecommons.org/licenses/by-nd/4.0/

many disciplines (i.e. law, semiotics, ergonomics, human-computer interaction, design, computer science, cognitive psychology, philosophy, behavioral economics, etc.) can offer valuable insight for the development and evaluation of data protection icons, it must be a goal of the regulators to find means and resources to carry out data protection education campaigns for European citizens. Indeed, the protection of personal data and privacy is one of the fundamental digital skills of the European Digital Framework for Citizens (DigComp) [12], which is the EU reference framework that provides the description of the competences needed in our digital age[13]. Besides, standardization initiatives can augment the knowledge of considerable quantities of population. Every segment of society should be possibly included and individuals of each European country should be reached. They should be diverse in terms of age, gender, educational background, profession, technical proficiency, legal knowledge, and privacy awareness. This is a very challenging and ambitious goal, but it is a necessary step to produce icons that can be safely used at the European level.

That said, it is impossible to produce an icon set that will be considered perfectly representative of data protection concepts, i.e. perfectly semantically transparent. Further testing can give important insights as for what concerns legibility, while alternatives for those symbols that scored worst can and should be sought. Indeed, as legal design presumes, there is no unique, fixed solution for a given challenge. Education on data protection matters can sensibly augment recognition rates. Nevertheless, very high rates of ease of recognition will never be reached for unfamiliar concepts or icons, until their widespread adoption will increase data subjects' familiarity. For this reason, standardization open to versioning is the path to follow: after one icon set has been publicly discussed and adopted, empirical data on its use in real-world scenarios should be gathered. Subsequent versions should then

[13] Among the many goals: "To protect personal data and privacy in digital environments. To understand how to share personally identifiable information while protecting self and others from dangers (e.g. fraud). To understand that digital services use a 'Privacy policy' to declare how personal data is used" [10, p. 7]

consider and integrate comments about the first version, together with the needs of evolving societies and regulations. Moreover, the evaluation criteria would benefit from an integration with qualitative methods: for future research, a mixed method approach [73] is therefore suggested.

DaPIS does not aspire to be the ultimate data protection icon set to be adopted at EU level: it is an experimentation of the possible methods of design and evaluation of graphical symbols for the data protection domain that aims to contribute to the scientific discussion and the evidence-based approach suggested by the Regulators and the WP29 (*see* Section 2.7). This is why, in the present contribution each design choice made about the icons is reported and examined thoroughly: so that public discussion and possible critiques can advance the icons' development.

6.8.2 Usability versus Precision

Throughout the phases of creation and evaluation of DaPIS, there was constant opposition between simplicity and preciseness of representation. Whereas the former is a fundamental feature to ensure usability and scalability of the visual elements to any dimension, the latter is important to convey the exact meaning of the corresponding concept. It is challenging to determine the extent to which a visual representation can be simplified without losing those necessary traits that contribute to convey its meaning. The risk of oversimplification is indeed one of the fears of legal experts (*see* Section 6.1.2). However, a trade-off must be struck, even if it is an arbitrary choice. Besides, although to a certain extent icons can convey unknown notions to data subjects, focus should be on other typologies of visual means that are more effective to reach this goal: pictograms, comics, infographics are obvious candidates, while animated gifs and videos can better convey movements and time sequences.

The two subsequent versions of the right to access and the right to data portability well illustrate this dichotomy and difference in scope: the first icon version supported a literal representation of the concept, while for usability reasons the subsequent versions were based on metaphors of less straightforward interpretation. In the literal visual transposition of the right to data portability (*see* Fig. 6.10) and the right of access, the movement of data between a user and the controller would result much clearer if animated, while the icon would be less crowded, since the actions would be shown consecutively. The literal visuals would be more informative and maybe more readily graspable. A similar stance can be adopted for the movement of arrows that attempt to convey the purposes.

6.8.3 Inherent Incompleteness

One of the goals of shaping icons on an ontological conceptualization of the GDPR is that of providing an icon set that could be employed in a variety of contexts. In other words, icon sets can be easily created for individual instances of privacy policies (i.e. for a specific service, for its specific data-handling practices, for the specific typologies of data processed) but then it would be hard to generalize the use of the same icons to a different context. Instead, the idea of basing the icons on an ontology has the goal of providing a visual language that is general enough to be adapted to any context, at least for those information items that the GDPR mandates.

For some concepts a finite set is given, for example there are exactly 6 types of legal bases and there are exactly 9 data subjects' rights. These concepts appear in every privacy policy addressed to Europeans and for this reason are readily codified in the ontology. Nevertheless, for other conceptual areas, completeness is more challenging to reach: for instance, the conceptualization made in the ontology about the purposes of processing could result to be too abstract to be of any practical use to the service provider. Indeed, data controllers often list the purposes of processing in a very detailed manner and these show great variety depending on the service domain. Specific visualizations could be recommendable, but then the icon set would become open-ended, e.g. for what concerns the data types, e.g. financial data, contact data, etc. For instance, for the Juro privacy policy [11], *ad hoc* icons were created. But due to the concrete nature of the concepts they represent, they can be easily represented through an example from the class they represent: for instance, a wallet or a money symbol for financial data.

That said, DaPIS' approach has deliberately aimed at standardization to constitute a replicable solution in as many contexts as possible, despite the shortcomings introduced by this choice. After attentive evaluation, it would probably be very useful to merge DaPIS with the Privacy Tech's icons [276] because the two sets are complementary. It is foreseeable that such a standardized approach would be employable by privacy policies editors that support standard structuring of these documents complemented with visual elements, like Signatu[14] or Iubenda[15].

6.8.4 Unsolved Problems

In conclusion, icons will not solve many problems outlined in Chapter 2: privacy policies as they are now fail to be informative and it is not solely a matter of information design. Their language is typically vague and ambiguous and the amount of information provided is usually excessive for data subjects, although suitable for those monitoring organizations' data practices such as regulators, supervisory authorities, and advocates [152]. Although with the transparency obligation many services sent updates to their customers claiming improved clarity and transparency, preliminary analysis of the privacy policies of the 14 world's main service providers [71] shows that around 11% of their clauses still contains unclear language, while around one third of the clauses are potentially unfair. Experience and empirical research indicate that users are desensitized by too many consent requests and that providing long and tedious privacy policies while the user is carrying out another task is only deemed as a nuisance. Behavioral insights can be used either to favor the data subjects or to deliberately obscure information and choices [148]. Finally, modern data processing activities have a high level of

[14] https://signatu.com/

[15] https://www.iubenda.com/en/

complexity. For all these reasons, it is unrealistic to expect that placing icons on privacy notices will alleviate this burden of explaining and understanding such intricacy. Icons alone will not attain the goal of transparency, but only constitute one of the visual devices that can try to solve the problems outlined in Section 2.2.

The GDPR should provide a strong incentive towards compliance with transparency principle, though. At the time of revision[16], the first financial penalty (50 millions euros) for lack of transparency has been imposed on Google [67]. Interdisciplinary studies can shed light on many overseen aspects of information transparency to empower data subjects to have more control over the flow of their data. New ways of communicating data practices are spreading, although they constitute only a minority [134]. In addition, further research about the icons' effectiveness must still be carried on and should benefit from more researchers and more perspectives. In the following some possible future directions of research are described.

[16] Early February 2019

Chapter 7

Future Direction and Open Problems

7.1 Future Directions

Notwithstanding the three evaluation studies carried out on DaPIS, a number of questions remain open and in need of further research. These matters are introduced and briefly discussed in the following sections and constitute recommendations for those that will want to continue to investigate the effectiveness of DaPIS or other data protection icon sets.

7.1.1 Effects of Training

Without the observation of users' progress over time, it is impossible to determine if DaPIS "works" or not. Since one of the main obstacles to ease of recognition is the lack of familiarity with the icon or with its referent (as emerged several times in Chapter 6), it should be expected that the effect of training increases recognition rates and determine easier recall. Thus, longitudinal studies should be preferred [109]. For example, comprehension tests should be designed in consequent and iterative steps: for instance, after the first comprehension test, a brief explanation of the symbol can be provided [299]. A second test, after a short amount of time, can be administered to the same subjects to evaluate whether their recognition rate increases over time and its false alarm rate decreases, as it is expected after a second exposure to the same icons. Learning ease can be determined also through more than two test reiterations and would arguably mirror more closely the actual users' sense-making of the icons in a privacy policy. Recording progress over time is more meaningful than a one-time only recognition test, especially within the view of affirming a standardized icon set.

For comparison, let's consider the learning path followed by a user while she develops confidence with an unfamiliar graphical user interface, for example that of a brand-new text editor: the user has previously developed mental models of the basic functions that any text editor software offers (e.g. print, save, new page, undo, etc.) and relies on them for orientation in the new interface. She will also look for familiar graphical symbols that stand for such functions. As she user gathers experience with the editor, she also develops a deeper understanding of the program's functions, while she also corrects false initial assumptions. Similarly, it is must be expected that, as the data protection icons' use spreads and data subjects gain familiarity with the symbols, they will also be able to recognize the icons more readily, until they will be able to understand them effortlessly even in isolation, i.e. without textual labels or explanations. This would be useful to decode icons as on IoT devices or during online transactions, e.g. while carrying out a different task.

7.1.1.1 Existing Testing Strategies

The procedure for testing comprehensibility described in ISO 9186-1:2014 [167], is devoted to symbols whose referents are familiar for the test participant, so that only the ability of the symbol to convey its meaning is actually assessed. This testing method is, however, not suitable for those symbols whose referent is unknown to the participants: in case of wrong association or inability to associate icon and referent, it is impossible to determine if the mismatch depends from the lack of knowledge about the concept or from the icon's scarce capacity to convey its meaning. As the first user testing revealed, this difference constitutes a serious concern to attain trustworthy and informative results.

A possibly helpful international evaluation framework for graphical symbols whose referents are unknown to the users is established by the international standard ISO 9186-3:2014 [168]. This standard introduces two consequent phases of testing: the first part aims to make users develop familiarity with the concepts (i.e. familiarity training), while the second part tests the comprehensibility of the graphical symbols (i.e. symbol referent association test).

The ISO's methodology envisages that participants learn a list of concepts and their definition. Then they are presented with the list of referents in random order and demonstrate their acquired knowledge by describing the meaning of the listed items. It is up to the researcher to determine whether the respondent shows adequate understanding of all the referents and, thus, the second phase can begin. In this part, for each symbol, 6 possible definitions among the ones in the list are provided to the respondents, while also the context where the symbol is meant to appear is described or shown in pictures.

This ISO constitutes a good methodological model to test icon sets like DaPIS, even if some modifications can be envisaged. The icon set consists of 37 elements: some are common or easy to understand, and therefore probably known to the general public (e.g. contract), others are very specific of a specialized knowledge (e.g. pseudonymization). However, given the fact that the icons should be understood by Europeans despite their backgrounds, it is not safe to assume *a priori* that people will be familiar with some but not with others: it is advisable that the familiarity training is performed on each and every concept, although more attention can be devoted to the more unfamiliar concepts. In the second place, the familiarization process described in the standard is more suitable for small lists of referents. Besides, more than one single judge should decide if the participant has internalized the concepts in order to be more objective.

7.1.1.2 A Proposal on Training Testing

Therefore, two adjustments to the ISO 9186-3:2014 methodology can be contemplated. Firstly, the training can be segmented in progressive steps, so that the participant is not confronted with a long list of concepts all at once. In fact, it would be ideal to facilitate the learning process by gradually building on previous knowledge. The progression could follow a conceptual categorization, while the concepts are presented in growing number and in increasing (presumed) complexity. Areas can be identified by the following questions:

1. Who? Data subject, controller;
2. Where? Storage inside EU, transfer outside EU;
3. How? Encryption, anonymization, pseudonymization, automated decision-making;
4. Why? Profiling, marketing, provision of the service, enhancement of the service, security, research, statistical purposes;
5. On which bases? Vital interest, public interest, legitimate interest, consent, contract, legal obligation;
6. Your rights? Right to be informed, to erasure, to rectification, access, data portability, withdraw consent, object to processing, restrict the processing, lodge a complaint to a supervisory authority.

Secondly, a more rigorous method to safely determine if the concepts have been acquired can be envisaged and repeated over time to follow the progression path of the learning process, until the user masters the entire list of concepts. A meticulous evaluation scale can be designed and the evaluation should be carried out by at least three independent judges or by a computer program that automatically assigns points to the answers according to strict and objective criteria. A similar procedure was designed for the first user study, but it was admittedly fuzzy and only one human judge was involved.

A possible alternative to the familiarity training of ordinary respondents would be to employ users that are already familiar with the referents: experts of data protection. This would spare the time and effort to create the above described assessment activity. Despite this advantage, the involvement of experts would not mirror the intended audience of DaPIS in its totality: it would only constitute a specialized subset of the audience, thus making the test results not generalizable. On the other hand, it can be argued that also the step of familiarization does not mimic real-world conditions: in fact, it is highly unrealistic that individuals confronted with the icons for the first time will learn the concepts behind them beforehand. Either way, the test conditions do not reflect real conditions - an issue that is widely debated for any lab test, for which however there exists no perfect and easy practicable solution. Real life allows for more learning through experience than lab studies, while lab tests, at least, allow for the isolation of variables and for the observation of specific phenomena.

The comprehensibility test would also need to be different from the ISO's specifications to account for the specific function of DaPIS. The data protection icons are not meant to appear in isolation on GUIs to signal the function of a software, as symbols on equipment or as warning symbols. They are rather meant to act as information-markers in a document in a complementary way with respect to the text. During the review of this dissertation, it was suggested to resort to the Semantic Inspection Method [81] to evaluate the icons and the interface design where the icons will appear.

7.1.2 Final Evaluation in Context

This section will propose some evaluation methods for the comprehensibility of DaPIS. After having tested the icons in isolation, as proposed in [223], a subsequent, necessary step to test the effectiveness of DaPIS would be to test the icons in a real context, as anticipated in Section 6.3). This means, for instance, that it is necessary to address questions about how users will make sense of the icons.

It is unrealistic to expect that data subjects will attentively read every line of a privacy policy every time that they encounter one. It is more realistic to expect that data subjects skim privacy policies to look for specific pieces of information in a limited time span, as research points out (*see* Section 2.2). Most of the times, the data subject are asked to agree to the conditions set forth in a privacy policy while executing a different task, e.g. while buying a flight ticket. In this case, the expectation that users will carefully read the privacy policy is simply unreasonable and the obligation to read the terms is experienced more as a nuisance than a legal safeguard. The WP29 recommends [30] to provide information through a link or on the same page where personal data is collected. Whereas the first solution would probably be a failure, the second would overload the page of excessive information. It is in such cases that a compact array of icons that summarizes the data practices could prove helpful: the first layer of a multi-layered approach. In this case, the icons would be interpreted as stand-alone elements to offer "in an easily visible, intelligible, and clearly legible manner a meaningful overview of the intended processing" (Art. 12.7 GDPR). Moreover, only icons showing potentially risky practices or perceived as such [197] (e.g. transfer outside EU, third party sharing) may be shown in order to attract the attention of the user, while leaving out the icons that are omnipresent. However, this would imply a one-size-fit-all-solution decided *a priori*, for which user-tailored communication could constitute an alternative (*see* Section 7.4).

Another realistic set where the interpretation of DaPIS can be tested would be an online, interactive interface where the icons complement the text and act as navigation cues. In this context and by reproducing realistic user tasks, an association task between symbol and referent would be meaningful if carried out as an information finding task. I.e., given a specific privacy policy, where would the user look for specific information items? On which icon would she click to open and expand the relevant section, provided that the policy is interactive and organized in meaningful paragraphs? Such a setting for the usability test would provide higher ecological validity to the study and probably determine higher icons' recognition rates (*see* Section 6.1).

Besides, researchers with expertise in visual perception should be involved in the study of how people apprehend, recognize and interpret these small visual elements. In the present study, enough time was allocated to a thorough examination of the icons, whilst in the real world the time might be limited to a glimpse over the icons. In that case, it would be meaningful to investigate if differences like white and black hands, or black and white users would be noticed.

7.1.3 Number of Icons

The icon set produced during the research described in these pages sums up to almost 40 icons. Although not all of the icons are expected to appear in a privacy policy at the same time and the ontology-based approach has hindered the proliferation of icons, research must be devoted to the cognitive overload that such a numerous icon set might cause, especially for first exposures. If the number of icons is too high to be easily apprhended, the risk of feeling overwhelmed might arise, which would be contrary to the very goals of DaPIS. Furthermore, whereas icons could prove beneficial in first trials because they can attract reader's attention or even curiosity thus acting against habituation effects, the effect of habituation over time must also be researched.

In this respect, it would be probably useful also to experiment whether the selection of a limited number of icons to be displayed would be more meaningful than the entire icon language. In particular, two scenarios can be envisioned. In the first one, only risky practices or practices that would have a significant effect on the individual are presented in a visual manner, such as the transfer of data outside of the EU and the existence of automated decision-making. The second scenario is shaped around the arising possibility of customization. Indeed, many studies show that expectations, needs, and fears around privacy depend on individual characteristics, thus they vary greatly. If, as proposed by the draft ePrivacy Regulation, browsers will directly manage data subjects' preferences, the possibility of being shown only the icons that matter to them could become a reality. Artificial intelligence could also play a role in this sense (*see* [145, 54, 55] and Section 7.4).

7.1.4 Discriminability

Another dimension that must still be scrutinized is the extent to which each icon is discernible from the others of the set, which is a crucial index for ease of recognition. Identifiability is indeed a relevant dimension: the icons are part of a set and the less they overlap in terms of similarity, the more they will be memorisable. In other words, their design should be sufficiently consistent to identify them as a family of icons, but also sufficiently distinctive to make each element easily distinguishable from the others. For instance, the difference between data subject and controller was not sufficiently relevant in the second user study, whereas the differences in the color of the user's silhouette (black, white, black-and-white) are admittedly rather subtle. Discriminability could be investigated with a hit rate task with multiple answers where all the icons under the same conceptual category (e.g. rights) or having similar meanings (e.g. consent, right to withdraw consent, etc) are displayed together.

7.1.5 Gamification

A longitudinal user study that addresses all the points listed in the previous sections would be complex and time-consuming, because the tasks would need to be distributed over a relevant span of time. An alternative, or at least a complementary way to financial contributions, to motivate participants to carry out the tasks and to continuously progress with the icons' learning is represented by gamification, which reflects the idea of using game design elements in non-game contexts [83]. In Section 5.2.8, a gamified experience was described as a possible path to beat the lack of motivation to read privacy policies. To foster motivation to provide long-term contributions, it can be useful and appropriate to rely on social-psychological processes like self-efficacy, group identification, and social approval, that stregthen the sense of competence and progress in the users [82]. Gamification is suitable in this context because it has a twofold purpose: firstly, offer users a fun and motivating experience and, secondly, leverage their playing time to tackle worthwhile endeavors [292].

The design of a gamified environment brings the advantage of encouraging its users to return on the platform to continue the progression and verify if their results increase (e.g. in terms of correct association icon - referent) as they gain experience over the icons and the underlying concepts. Badges and points are typical game mechanics that make the progression tangible for the players and increase their motivation to continue playing and achieve better results.

7.1.5.1 Tasks

Within this view, in such a complex but integral experience the users would have the possibility to:

1. progressively familiarize with the referents (i.e. familiarity training);

2. progressively familiarize with the icons (i.e. familiarity training);

3. progress in the learning of the right association between icon and referent (i.e. comprehensibility);

4. learn to discriminate among icons pertaining (or not) to the same conceptual category (i.e. discernibility);

5. learn to associate textual parts of a privacy policy with the corresponding meaning.

For what concerns comprehensibility, a blind privacy policy with icons can be presented to the user, who is asked to match each icon with the corresponding section of the text. The number of correct associations can be reasonably expected to increase as users progress, while errors should decrease. The error count would be useful to determine which icons are less readily associated to the corresponding meaning. A critical threshold could be elected to determine if alternatives should be sought for icons performing particularly badly. A similar task can be envisioned to establish how easy it is to distinguish one icon from the other: a marked up privacy policy is provided to the user who has the task to match each section with the corresponding symbol taken from the entire pool of icons.

7.1.5.2 Players' Profiles

A gamified environment offers the possibility to differentiate among the different profiles of the players. Players can be novices, tech-savvy people, or legal experts and be, thus, assigned tasks that depend on their level of expertise. For example, the task of annotation presumes a comprehensive and sound knowledge of matters related to data protection: the profile of legal experts is much more suitable for such an activity. Legal experts might also need a shorter familiarity training compared to others and might perform the icon interpretation task with less effort. Different success levels must be expected and the performances across groups with different characteristics can be thereby compared, as it should be done in any user study.

7.1.6 Usability and User Experience

Once that the icon set has been evaluated in all the above listed dimensions, it is necessary to gauge the usability of the visualized privacy policy as an artifact. As defined in the Section 4.2, usability is the "effectiveness, efficiency, and satisfaction with which specified users achieve specified goals in particular environments" [166]. In the 'environment' of a privacy policy, the dimensions that should be considered and measured are:

1. effectiveness: achieve the goal, i.e. find relevant information in the document

2. efficiency: achieve the goal with low effort, i.e. find relevant information in the document easily and quickly

3. satisfaction: achieve the goal with positive feelings, i.e. have a good user experience with the document

For what concerns the first two dimensions, user research on contract visualization (*see* Section 4.3.4) can provide a useful and replicable methodology. Accuracy and speed of comprehension are relevant criteria: it must be determined if a privacy policy with icons increases accuracy of answers and decreases speed of comprehension[1] with respect to a traditional text-only privacy policy with the same contents (i.e. the control condition). This is what is called an A/B test.

In addition, it would be necessary to consider possible effects of information architecture on accuracy and speed of answers. In this case, three experimental conditions should be envisaged to determine if, for instance, information architecture *per se* is more or less effective than the icons: the users' performance on the text-only privacy policy, structured privacy policy, and visual privacy policy should be thus be assessed. In any case, the hypothesis would test if the visual privacy policy determines more positive outcomes in terms of ease of information finding and time consumption to carry out the task compared to the text-only privacy policy.

Finally, the user experience should also be considered to determine the users' perception and interaction with the visual privacy policy. The extreme length of traditional privacy policies makes the data subject feel helpless and frustrated, whereas a comprehensible and navigable text can trigger positive

[1] Although accuracy and speed are classical usability measures to determine efficiency, doubts can be harbored for the appropriateness of accuracy in the specific case of legal icons. Since icons are simple elements and convey meanings also in terms of metaphors, it would be critical to assess whether, despite these features, they can contribute to a more accurate comprehension of the legal terms.

emotions, such as satisfaction [229]. There exists several scales to attest the user experience: for instance, ease of use, satisfaction questionnaires, expectation measures, etc. (*see* for more details [19] and [267]).

7.1.7 Costs of User Testing

As argued in the previous section, gamification can be a helpful resource to address many issues. It is recognized [14], and confirmed by the experience described in these pages, that the main drawback of user quantitative testing is its costs, in terms of human and financial resources. Moreover, it takes time to find the study participants, hypothesize and prepare the testing materials, gather and analyze the data. The situation even worsens if the user testing is iterative. Nevertheless, user testing at any stage must be considered as an investment, since it can spare future costs of implementation of a non-functional solution (*see* Section 4.2). Even only a few users can point out the main pitfalls of a certain design. Qualitative testing options need also to be carefully examined.

However, there is value in the standardization of the process of generation of visual elements, especially if this approach is combined with ICTs. If the legal drafting relies on standardized procedures, and as such, is supported by technical tools (such as editors) that recognize common patterns and suggest desirable formulations, templates and structuring, it is foreseeable to extend this application to the visual domain and even to integrate visual patterns (such as icons) into the user software. Furthermore, the generation of visual elements to serve legal purposes might even become automated, as the subsequent search and retrieval activity, like the approach presented in this dissertation suggests. The automation will thus reduce the high encoding costs sustained at the beginning, and will also cause reduced costs of transmission, retrieval, and de-coding [42].

This last point introduces the next sections that aim to open a discussion upon critical points that emerged during the design and consequent evaluation of DaPIS. Participatory methods for legal icons' design constitute an innovative experience and have raised some open questions that have not found an answer yet.

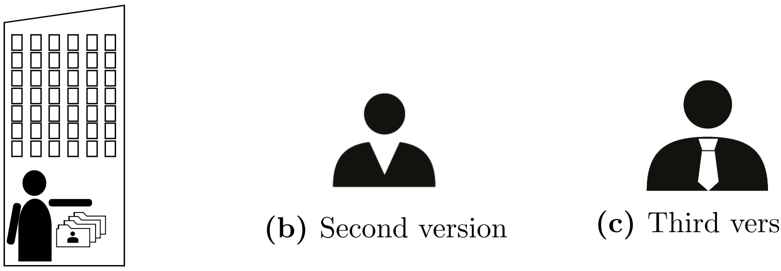
7.2 Open Questions

7.2.1 Reconciling Different Mental Models and Priorities

The collaboration among professionals with different backgrounds in participatory design can be transformed in a profoundly fruitful exchange. It lowers the chances of personal bias derived by one's own mental models and leverage on the unique values, skills, tools, and knowledge that each background contributes with [253]. To produce icons for data protection, the empirical experience described in these pages suggests that each participant (i.e. legal expert, designer, computer scientist, but also layperson) has different views, mental models, memories, and intuitions, but that all can contribute in the transformation of complex legal-technical concepts into effective graphical representations.

However, this can also show some tensions, as the many cited cases of fierce opposition between simplicity and precision of representation demonstrate. As an example, let's consider the discussion that raged around the concept of "data controller" during one of the workshops. The first point of debate revolved around the question of whether the controller should be represented as a legal person or a natural person. Earlier examples of iconography for this concept tend to represent it as a tall building, that evokes a corporation or organization. Early trials of user testing confirmed that individuals can recognize the concept behind such visualization, provided that some context is given.

Yet, the role of data controller with legal responsibility is rather an individual that represents the organization and it is mentioned in the privacy policy as such. Representing this role as an individual was considered misleading by some legal experts, whereas others argued that even natural persons (e.g. privates) can be data controllers. During the first workshop, it was therefore decided to represent this agent as a combination of a tall building and a person in the building (*see* Fig. 7.1a). However, despite its appropriateness with respect to the legal definition, this visualization is too complicated to be rendered in small sizes. Indeed, the first user study confirmed that the icon showed excessive complexity. During the second workshop, the discussion was enriched by additional considerations, for instance by the reflection that the figure of controller should also be combined with other building blocks, for instance in the contract icon. Hence, the icon for controller must be as simple as possible: its representation as a business man was deemed a pragmatic, acceptable compromise, even if not perfectly appropriate (*see* Fig. 7.1b). Excessive simplicity in the icon design, however, resulted in a low level of discernibility of the controller icon from the data subject icon, as the second user study revealed. A third elaboration of the icon that emphasizes its distinctive elements, if compared to the user, had to be finally proposed (*see* Fig. 7.1c). Ultimately, a simplified version of the building can be reconsidered for its easy recognizability.



(a) First version

(b) Second version

(c) Third version

Figure 7.1: The evolution of the icon for the concept of controller

7.2.2 Definition-Dependent Visualizations?

An additional, subtle problem is given by the wording in which some concepts are expressed. For example, the GDPR defines the third party as "a natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data" (Art. 4.10). The definition is then given in a negative rather than affirmative form: it defines what the third party is not, while omitting to define what the third party is. A possible reason is the fact that the concept is vast and enough space for interpretation is left. This circumstance, however, hinders the creation of an appropriate visualization. The adopted solution considers a practical and functional stance: the typical context where third parties appear in a privacy-related communication mentions the sharing of personal data with third parties. Thus, from an ontological point of view, the focus is shifted on the action of sharing with third parties, rather than on the role of third party.

7.2.3 Language-Dependent Visualizations?

Another question concerns the possibility that the language in which a concept is described influences its consequent visualization. Indeed, a commonly used method to spark the design process is the brainstorming activity around a concept, which is necessary to introduce new ideas when an established visual vocabulary is missing. However, the visualizations that words evoke can be language-dependent.

For instance, the first workshop described in Section 6.2.2.1 was held in English, while the consequent ones in Italian. In both, the concept of "supervisory authority" was thoroughly analyzed to produce an icon for it. But, whereas in English "authority" evokes power, and even command, while "supervision" can be associated to a police activity, the Italian correspondent "garante della protezione dei dati" rather evokes the idea of a person guaranteeing the data subjects' rights and a situation of equality where all interests are considered and balanced. The implications are not irrelevant, since the two translations suggest different perceptions of the supervisory authority. The visual representations proposed for the concept in the different workshops reflects indeed the influence of language. In the end, a more neutral representation was chosen: that of a person sitting at a large desk, suggesting the idea of a person that scrutinizes wrong doings. Other doubts derive, for example, from the representation of legal basis that was coined on the verbal definition: the basis is represented as a column, while the legal is represented as a gavel. If in other languages the same concept is linguistically realized in a different way, then it is doubtful whether the metaphor can be easily recognized.

7.2.4 Culture-Dependent Visualizations?

As the Primelife experience pointed out [155], the available visual vocabulary of an individual is inescapable from her cultural background. For instance, during the first workshop described in these pages, it was decided to represent data processing operations as a simplified narrative from left to right, where the personal data folder undergoes the processing, and, as a result, is transformed into a different kind of data (*see* Fig. 6.7). However, the left-to-right narrative replicates the direction of writing followed in Western countries and could result rather innatural for populations with different writing traditions (e.g. from right to left or from top to bottom). Even the vertical, as opposed to horizontal, direction can suggest some semantic information. For instance, in the second study a participant (P13) describes the icon for the supervisory authority as "man in an office (probably representing a subject of supreme authority) and below him a folder with a man that is inferior to the first subject". The intention of the designer was indeed that of giving a certain predominance to the authority also by placing it on top of the data folder. Although the left-to-right and top-to-bottom directions would nevertheless be suitable for an EU context, further research and considerations on the topic are needed if the icons need to reach global acceptance.

Also the different meanings of colors can stimulate a similar discussion. In the present research, it was intentionally decided to exclusively use black color and blank spaces, that appear white on white backgrounds, as good practice[2] suggests. However, even these two colors assume meanings, which are culture-dependent (i.e. Western versus Eastern tradition). A concrete example is represented, again, by the supervisory authority: the preponderance of black in the first version of the icon was interpreted by many test's participants in a negative sense, as they wondered if the figure was a "villain" or a "cyber-attacker". Evidently, even the colors of the icon suggested a different meaning than the one it was supposed to. On the basis of these comments, in the second version the authority was designed with a prevalence of white. Finally, although the use of black and white was deliberate, research in this sense should be brought forward to determine if colored icons would be more easily recognizable than their black-and-white counterparts because more realistic.

7.2.5 Time-Dependent Visualizations?

Another topic of debate emerged during the workshops is the time-dependency of certain visual choices. The visualization of the legal basis of consent, for instance, generated a lively discussion in our second workshop. Consent as legal basis, as expressly defined by the GDPR, should be a free choice between the act of giving consent and that of withholding it, indicating the data subject's agreement or disagreement with the processing of her data. In an online context, consent is mostly given through the click on a button or the ticking of a box. Initally, a mouse cursor between an acceptance button (a thick) and a refusal one (an x) seemed an appropriate metaphor to signify the possibility of free choice. Although it can be argued that this would have simply constituted an exemplifying icon for the many manners in which consent can be given, the mouse cursor was deemed a far too technology-dependent, therefore time-dependent, element. Doubts were raised on the understandibility of this symbol over generations (e.g. older versus younger

[2] *see* e.g. the Nounproject's technical guidelines `https://thenounproject.com/handbook/create/#technical_guidelines` and ISO's recommendations [168].

users) and over time (i.e. consent can be given by a physical signature, a click with a mouse, but also a gestural interaction, voice, and, in the future, iris scanning or similar).

However, there are other cases of icons whose meanings and whose representations have fossilized over time, while the actual referent has changed. A classical example is constitued by the save function represented as a floppy disk: although young generations do not know the object, they have learnt to recognize the icon and its meaning.

7.2.6 Priming to Practical Examples

Practical and concrete examples can be very valuable when designers are trying to visualize a certain concept with whom they are not familiar. In our workshops, it was considered necessary to provide simplified definitions and examples to the participants. Sometimes icons depict representative individuals of the class of entities they represent (i.e. exemplar icons: the fork and knife to signify a restaurant) because it is more practical and easier to do so, than to try to represent the above, sometimes abstract class. Use-cases, exemplification and story-telling are also one step of the method used to build (legal) ontologies because they provide concreteness to concepts that would otherwise be too abstract to be easily grasped and modelled. Indeed, as the first experiment shows, scenarios have been more appreciated and have been more easily related to the subject's own experience than the simplified definitions of the same concepts, despite the easier language employed.

However, providing examples to stimulate the process of creation could also have another effect, namely that of influencing the object that the designer will choose for the visualization. In this case, there could intervene what is called a priming effect: providing certain examples activate certain associated memories, which then might influence the icon creation process. For instance, the provision of an example related to medical research to explain the concept of scientific purpose might have influenced its visualization as a microscope. But only psychological research can provide answers to this point.

7.2.7 Embodied Cognition

Another relevant direction of research, that might need to be considered in the future, is the study of legal concepts from a psychological perspective and, specifically, that of embodied cognition. Some scholars [256] raise the question about the possibility of linking how legal artifacts are built and conceptualized with the way in which natural and physical phenomena are conceptualized: embodied and grounded views maintain that even abstract concepts are grounded in the physical interaction with environment. Although these types of studies are very innovative in their nature, hence the number of concepts investigated is still low, they could possibly offer an additional frame to guide the creation of legal icons that resonate with the common characteristics of human experience. For instance, the concept of authority, can be traced to an image schema that represents something that has a causal effect on other bodies, whilst contract can be grounded into an image schema of contact and transmission between individuals [256].

There is also a growing body of research that suggests the study of abstract concepts in a more fine-grained manner than it is usually done, instead of considering them as an undifferentiated, unique class [121]. According to such a categorization, it could therefore be researched whether not only the level of abstractness of icons is influential in their ease of comprehension and recognition, but even if differences can be noticed according to the characteristics of the class to which abstract concepts belong.

7.2.8 Digital versus Physical Apprehension

The GDPR emphasizes that special attention should be devoted to any communication addressed to children. Not only child-friendly language should be employed, but the WP29 [30] also encourages experimentation through other mediums: “[w]here transparency information is directed at children specifically, controllers should consider what types of measures may be particularly accessible to children (e.g. these might be comics/ cartoons, pictograms, animations, etc. amongst other measures)" (p. 12). With respect to other generations, digital natives experience the reality differently, process info-graphics speedily, and learn differently, while their brain presents neruological differences as a response to digital experiences [65]: "[d]igital natives can process images faster and compartmentalize information more efficiently, yet they often lack long-term memory functionality and the cognitive functioning required for critical thinking and imagination" [65, p. 4].

Research also shows that individuals make sense of information differently on screen than on paper (e.g. [195, 76], even if contrasting results exist [244, 157, 250]. The digital revolution has been so disruptive, but at the same time is so recent, that any study warns that it is too early to speak of definitive results. It is undeniable, though, that the reading activity has dramatically changed, as individuals process increasing amounts of information at a faster pace. Nevertheless, digital reading, if compared to paper reading, also causes differences in levels of attentiveness, long-term memory function, depth of comprehension, and active reflection upon what is read (for a comprehensive analysis, *see* [65]).

The predominance of visual culture and of a visual vocabulary derived from electronic devices also strikingly emerged during the user studies: the participants made reference to the icons found in smartphones, for instance the symbol of gears to indicate settings, the edit symbol of the Facebook app, or the sharing symbol on many social media. However, it is not said that these observations can be generalizable to older generations or to a less tech-savvy population.

Hence, it is reasonable to expect that the design of a successful privacy-related communication should pay attention to the emerging characteristics, but also the issues, of a digital society. Visual communication seems particularly profitable because it permits to convey more information in a limited time and space, while at least digital natives seem to be more literate about image sense-making than text sense-making. Besides, visual cues complement textual cues, and are processed by different parts of the brain (*see* Section 4.3): one type of information acquisition can complement and reinforce the other.

7.2.9 Reading versus Seeing

This last argument introduces another possible future direction of research that addresses the question about differences in the sense-making activity: could the legal message be internalized differently if it is seen from when it is read? The law is traditionally written in the form of a discourse. It can be argued that the brain processes information differently if it reads, sees or hears it. Whereas reading is an activity that happens over time and gives individuals the possibility to pause and reflect, showing a message through an icon is rather an immediate apprehension. Although the icons are meant to complement text rather than replace it, it can not be guaranteed that individuals will always make sense of them in combination with their textual explanation. This can be particularly important in the case of digital natives, people with low literacy, non-natives, or people that are language-impaired (e.g. dyslexia, etc.). If the icons must give a meaningful overview of the intended processing, as the GDPR indicates, this means that at a first glance, data subjects should be able to understand what happens to their data, given the presence (and disposition) of certain icons in the privacy policy. Further research about the inferences that individuals draw from a visualized privacy policy should be therefore encouraged.

7.2.10 Predominance of the Text over Visuals?

Even if data subjects are expected to carefully read the privacy policy, research shows that this does not happen: they rather make assumptions based on heuristics. Icons can attract attention and quickly convey meanings, but there is the risk that for some people they will act as substitute for text. This eventuality also introduces an additional problematic question: which has preponderant legal value, the text or the image? The adherence between icon and underlying message is important in case of a dispute. The methodology used to design the icons (i.e. participatory design) and the iterative evaluation was expressly meant to produce icons that adhere as closely as possible to their meaning. Ideally, in a pro-active perspective, the visualizations will make the meaning of the text so clear that chances of misinterpretation will lower (*see* Chapter 4). However, the discussion in these pages has made clear that it is nearly impossible to induce rigidly deterministic interpretation results in addressees of the message, because too many variables, among which context and individual characteristics, enter the picture and influence the interpretation activity. It has been suggested elsewhere [236, 233], that visualized legal terms will bear a relationship with textual terms similar to legal texts translated in different languages: one form prevails on the other in the case of a dispute. A research area around visual jurisprudence is arising (*see* [46, 201, 209, 245, 268]), which represents the appropriate opportunity to develop arguments around the issue.

7.3 Icons' Context of Use

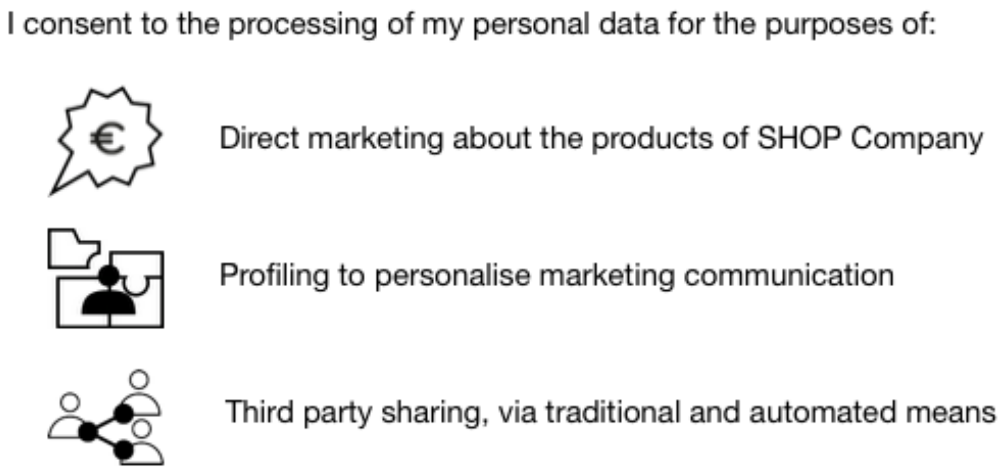
Several times during this dissertation, it was argued that considering the icons as stand-alone elements, i.e. outside of their actual context of use, can thwart their interpretation. In fact, DaPIS was designed with explicit attention to the role of icons as functional elements in a specific contextual situation: icons as information markers for lengthy privacy policies. An illustrative skeleton of a privacy policy that combines structured layout and icons as aids to navigation is displayed in Annex A.

However, icons can also be useful information markers in a multi-layered approach, where, as the WP29 suggests [30, p. 18], the first layer should "include the details of the purposes of processing, the identity of controller and a description of the data subject's rights [... It] should also contain information on the processing which has the most impact on the data subject and processing which could surprise them. [...] Therefore, the data subject should be able to understand from information contained in the first layer/ modality what the consequences of the processing in question will be for the data subject". The display of the first layer of written information is helpful especially in those occasions where the user is carrying out a task (e.g. a ticket purchase) and the reading of the privacy terms would be regarded as a nuisance. Since the caption "I have read and accept the terms" is omnipresent, but completely disregarded, a handy summary through icons could constitute a possible solution. An example is provided in Annex B. As familiarity with the symbols increases, it is foreseeable that it will become possible to display the icons with a minimal amount of text or even without any text - this could be particularly profitable on IoT devices that only have small screens or none at all.

Similarly, icons can be employed in consent forms, not only to quickly convey the purposes for which consent is required (as in Fig. 7.2), but also as actionable elements to design new, more meaningful, interactive experiences. Icons that function as buttons to actively and deliberately show user's consent to certain practices or in drag-and-drop agreements [108] can be experimented. Figures 3.5, 3.6, and 3.7 some examples of the use of icons combined with machine-readable logical structures extracted from the legal terms was also provided.

7.4 Customized Visualizations and Disclosures

The results of the tests described in these pages, as well as the unsolved question about the threshold of acceptability of a certain icon, lead to the question on whether or not the standardization of the visual elements can be the ultimate solution to the inevitable variance of interpretation. Without any doubt, as data subjects grow familiar with data protection concepts and icons, the chances of misinterpretation of the visual elements decrease (*see*



I consent to the processing of my personal data for the purposes of:

Direct marketing about the products of SHOP Company

Profiling to personalise marketing communication

Third party sharing, via traditional and automated means

Figure 7.2: A possible visualization of a prototypical consent form, where consent for the purposes of marketing, profiling and thrid party sharing is asked.

Section 7.1.1). Gaining experience, especially if through a process of trial-and-error creates a mental model of such experience in the individual, which can resort on it to interpret and react to future, similar events.

However, if what is sought is the certitude that the right message gets across, than this will most probably prove impossible. Besides, it might be the wrong goal to pursue. It is unreasonable to expect that individual differences (for instance in terms of age or experience) do not play a relevant role in the interpretation of any kind of communication, let alone visual communication. Too many factors prevent to deterministically reach certain results: from individual features of the people involved in the communicative exchange, till ungovernable contextual dimensions. Therefore, it cannot be expected that the interpretative process operates as an algorithm, where given a certain input, definite and foreseeable outcomes are produced.

Nevertheless, strategies to guide the process of interpretation towards desirable results can be employed. In the first Chapter, some of these mechanisms were named, among which user-centeredness (similarly to what suggested by the Article 29 WP [28]). At the preliminary stage of this research, we proposed user-tailored data protection icons, customized on the basis of the user's characteristics (e.g. age), to deal with the impossibility of total efficacy of such information [255]. Similarly, in [254], we proposed tailored consent forms that use personalized consent requests and visualizations. Although such lines of research were promising, they have not been (yet) implemented. Two main valid criticisms can be prompted by the idea of user-tailored communication. Firstly, personalized disclosures would be based on profiling and automated decision-making, which would cause privacy concerns and would have to be subject to the GDPR. Besides, law provides general and abstract rules, that have to account for a number of individual cases. However, it is exactly this abstract and impersonal dimension that causes disclosures to be unhelpful and meaningless for the data subject. In this light, the granularity offered by tailored disclosure would compensate the shortcomings of general rules and minimize regulatory errors (*see* [55] and [139]) to generate disclosures that enforce the right to be informed in an effective manner: "smart disclosures" [54].

Smart or personalized disclosures as regulatory tools have recently entered the scientific discussion in data protection law, as in other areas of the law: instead of providing standardized, one-size-fits-all disclosures that do not consider informational needs of the individual, we can envision the design of tailored disclosures that take into account individual characteristics (e.g. in terms of interests, concerns, or expectations) and show the advantage of providing relevant information, without the risk of information overload. Busch [55] invokes granularity and personalization based on big data analytic and profiling of the individual as solution to over- or under-inclusive norms. The argument behind such ideas rests on the fact that research has shown that simplification can not be the ultimate solution to privacy policies: as also recalled several times in this dissertation, simplification risks too easily to become oversimplification and in some cases can not account for the complexity of data collecting and processing. This argument must not be understood as a blanket rejection of simplification: the fact that the language of the law is in most cases overly complex and not addressed to an average person is an undeniable fact, that can be mended. Nevertheless, simplification can not be understood as the only viable solution, also because it is based itself on standardization. Instead, customized disclosures could prominently tackle the multiple privacy preferences of different data subjects (e.g. in terms of type of data collected e.g. geolocation or purposes e.g. marketing).

Profiling based on past user's behavior allows predictions about her privacy preferences, which has the advantage that users do not have to constantly make decisions, thus avoiding consent fatigue. Combined to the growing number of interconnected IoT devices that capture every intimate aspect of our life, the number of notifications and consent requests about personal data collection and processing is simply unbearable. Recent findings show that preferences about privacy are diverse and context-dependent, but can be predicted with accuracy by observing people's behaviors in a few scenarios [210]. Privacy preferences can be thus automatically adjusted and set as defaults, without the need of constant individual's intervention. Similarly, notices can be adapted to the person's behavioural patterns and privacy preferences, and to contextual elements. Such an approach becomes even more important in an interconnected world, where most IoT devices have no interface to communicate with the data subject only have such small screens that the display of notices becomes even more challenging than on any other device [8].

Artificial intelligence can also assign icons based on completely automated analysis of privacy policies, as shown by Polisis [145], and can even visually flag if a practice is occurring or not according to the categorization given by CREATe. Similarly, Claudette [71] is trained to automatically detect unfair clauses under the GPDR in terms of information completeness, transparency and lawfulness. Nevertheless, there is no artificial intelligence yet that can be totally trusted for the analysis and interpretation of legal language without any human intervention.

In the future, it is also possible to envision the combination of smart disclosures with icons that are customized for the intended audience. For instance, the workshops described earlier suggested the idea that average users and legal experts might want, and need, different graphical symbols for the concept of “data controller”. A common solution might be challenging to be found, whilst customization could be more feasible in terms of outcomes. Whereas the first goal is standardization of information provision, even in terms of icons, individuals are more and more used to receive content in a personalized manner that considers their profile, interests, attitudes, etc. This is also a line of research that is starting to be explored in the context of privacy-related information provision. This would, however, most certainly pose some legal problem. For instance, who would operate the selection and on which criteria? If a different version of the same privacy policy or different visuals were provided according to individual’s profile, which one would be legally binding? Would the focus on interests-based terms cause another sort of filter bubble, where users would be informed only about what has been deemed relevant for them? Such discussion was started during one of the workshops held in the context of the present research (*see* [142]), but has not reached definitive answers. In the meantime, a viable solution is that of offering uniform information, but also providing filters and options for customization to give to data subjects the possibility to manually select the visualization that works better for them, for their mental models, and for their memory.

In conclusion, it is fundamental to highlight and clarify a point that seems too often overlooked in the discussion about icons and transparency mechanisms: privacy policies are not supposed to convey new knowledge to the reader, as if they were educative instruments. Education about data protection can be reached through different means and in dedicated spaces, but it should be a critical component of everybody’s modern digital life. Research about privacy communication is intensifying and innovation is spreading across businesses, due to the GDPR’s transparency obligation. It is to be hoped that such experimentation continues and gains widespread acceptance.

Chapter 8

Conclusions

The present work has generally revolved around the key role that design, and legal design in particular, can play to effectively apply the principles and enforce the provisions of the General Data Protection Regulation, the EU legislation that has come into force in May 25, 2018. Chapter 2 has analyzed the scenario where the research described in these pages has developed. Firstly, the historical origins and the reasons for the introduction of the information paradigm in EU data protection law have been retraced. The information paradigm originated in a time when the advent of digitization and concentration of data in the hands of a few entities caused a general perception of loss of control over the personal information, while at the same time individuals were mostly unaware of such data gathering. Hence, in order to rebalance the information asymmetry between individuals and entities, regulators introduced mandated disclosures about personal data collection, combined with the instrument of consent to enhance individuals' control over the processing. Notwithstanding the massive data revolution happening in the contemporary digital age, where enormous quantities of personal data are gathered and analyzed ubiquitously, the European legislation is still based on the paradigm of transparency and choice, realized by the tools of disclosures (i.e. privacy policies) and consent requests.

Despite the value of such regulatory instruments, an extensive body of literature from various disciplines has demonstrated that there exist several hurdles to an effective implementation of such mechanism in practice. Individuals tend to disregard privacy terms and to experience consent requests as a nuisance, instead of leveraging them as instruments to exercise their rights to protect their private sphere and to express their privacy preferences. Our analysis has identified general tendencies that counteract the willingness and the actual ability to read and understand privacy policies. The language of privacy communication is usually complex and legalistic, while resulting at the same time vague, and it is not tailored to its supposed addressees, the data subjects. Furthermore, privacy policies rarely offer information architecture and mostly appear as a wall of text, discouraging the individual from engaging with them and hindering content navigation and strategic reading. This limitation also affects the possibility of comparison across different providers offering similar services, thus impeding the possibility to assess and critically compare risks and benefits of the processing. The excessive length of the texts, multiplied by the huge number of privacy policies that an individual is expected to read, makes it impossible even for well-motivated data subjects to be informed about the use of their personal data. The fact that relevant information is usually proposed at set up time, for instance when individuals sign up for a service, and not at the time when privacy-related decisions are taken, also constitutes an hindrance to privacy-conscious behaviors. Lastly, data subjects are treated as competent overseers of their privacy, while in fact they do not have enough expertise and knowledge to understand and assess the consequences of their disclosure attitudes. This aspect is also linked to the extreme complexity of nowadays' data processing, that contributes to make the explanations about data processing even more wordy and complex. Since information is the necessary precondition for informed consent, the shortcomings of mandated disclosure also exercise a negative influence on the possibility of provision of a freely given informed consent.

All these obstacles to effective communication have been esplored in combination with the evidence derived from behavioral studies that individuals are limited by bounded rationality, whereas the law presumes rational and attentive decision-makers. Therefore, Chapter 2 has also analyzed the hurdles to rational decision-making that data subjects experience not only in the privacy sphere, but in any domain of life: individuals base their choices on rule of thumbs (i.e. heuristics) and cognitive biases, which are systematic deviations from behaviors postulated by rational economic theory. For instance, salience of some information items over other items influence people's understanding (i.e. an effect known as framing), while individuals tend to stick to the *status quo* (i.e. the inertia bias). This is why the debate around nudges has been introduced: nudges are changes in choice architecture that leverage individuals' cognitive biases to encourage (or conversely discourage) certain behaviors. In the privacy domain, nudges like privacy-friendly consent defaults can effectively support data protection, while structure and framing of information can ease data subject's understanding of legal terms.

Whereas the several shortcomings outlined above have brought some scholars to call for a complete abandon of mandated disclosure, we have proposed to analyze how changes in information architecture can fruitfully help data subjects to cope with the complexity of data processing. We have thus described how the General Data Protection Regulation (GDPR) has taken into consideration much criticism towards the transparency and choice paradigm and has provided solutions in its provisions: data protection by design and by default constitute one of the main novel principles introduced by this piece of legislation, while consent must be signified by a clear, affirmative action and cannot be based on the inactivity of the user. In addition, the very concept of transparency is revolutionized: not only, under the GDPR, it becomes an obligation imposed on data controllers, but also it reflects an unprecedented attention to the quality and factual comprehensibility of information that must be offered to data subjects. In other words, privacy policies that focus on merely covering data controllers' liabilities are explicitly banned. "The concept of transparency in the GDPR is user-centric rather than legalistic" has argued the WP29: the characteristics of the intended audience and of human cognition must be taken into account to provide effective information of privacy communication. Not only language and information architecture receive an unrivaled consideration in legal communication, but also visual means are explicitly suggested to comply with the principle of transparency. In particular, Article 12 of the GDPR establishes the provision of information to data subjects in combination with machine-readable, standardised icons. This suggested measure has set the foundation for the research described in this work and has identified two combined directions of investigation: on the one hand, one line of research concerning the technologies that allow machines to understand the semantic meaning extracted from legal documents and, on the other hand, one line of research grounded in a human-centered approach to the law and to legal information.

The first direction has been explored in Chapter 3 and was based on the assumption that the transformation of legal content in a machine-readable form can also be leveraged to semi-automatically display human-readable information, in terms of structure and visualizations. The chapter has thus described standard formats for the management of legal content and its machine-interpretable description in terms of structure, semantics and rules: Akoma Ntoso, OWL, and LegalRuleML. The Akoma Ntoso XML schema provides a vocabulary to capture structural and semantic elements of legal documents. Moreover, it provides mechanisms for the reference to external ontologies, that are semantic resources that formally represent (a domain of) reality to enable the sharing of information and knowledge about it. Finally legal rules (e.g. permissions, prohibitions, etc.) can be modeled through the LegalRuleML XML-based rule interchange language. In this chapter, several examples were cited to show how such machine-readable information can be visualized: whereas visualized legal rules can, for instance, clearly show the consequences of providing or retaining consent, Akoma Ntoso provides structural elements that can be leveraged to enhance layout and information architecture. Not only: its semantic tags and the corresponding ontological concepts can be used semi-automatically display the visualization of data protection concepts, as the GDPR suggests. This is why the last part of the chapter was dedicated to the description of the design of PrOnto, a GDPR-centered privacy and data protection ontology, that has served as semantic foundation and conceptual organization for the design of DaPIS, the icon set at the center of this research. PrOnto has been organized in conceptual modules: i) data (e.g. personal data); ii) agents and roles (e.g. data subject, controller); iii) data processing operations (e.g. anonymization, encryption); iv) processing purposes (e.g. marketing, profiling) and legal bases (e.g. contract, legal obligation); v) legal rules and deontic operators (e.g. data subjects' rights).

Chapter 4 has described the integration of this technological standpoint with a human-centered approach to law and legal information: legal design. Before describing methods and tools of this emerging discipline, we have introduced the notions of legal literacy that presumes that understanding is not a simple action of acquisition of information. On the contrary, it entails the ability of decision-making (e.g. informed consent) and action (e.g. the exercise of a right) based on such information. Similarly, document literacy goes beyond reading to include searching documents to answer one's own questions and determine the relevance of information. Within this view, legal documents such as privacy policies must be considered not only as mere containers of legal information, but as effective tools structured around users that should support the satisfaction of their informational needs.

This stance is shared by legal design, an emerging discipline that has been described as "the application of human-centered design to the world of law, to make legal systems and services more human-centered, usable, and satisfying" [141, Chap. 1]. Again, user- or human-centeredness is a key notion and a driving assumption to carry out user research in the legal field. Legal design offers tools to bridge the gap between the theoretical assumptions of the law and the concrete individual's behaviors illustrated in Chapter 2. We have therefore illustrated the connection between this discipline, the plain language movement and the proactive approach to law. We have also described the usefulness of empirically-based research methods: the end-users of a certain legal artifact, such as a legal document, can be involved in every phase of the design cycle, from the analysis of their needs, to the development of solutions, till their evaluation. Such a partecipatory design stance has been adopted for the design and evaluation of DaPIS, as outlined later.

Indeed, another pillar of legal design is the reliance on visual methods: visualizations have entered the realm of written law and evidence-based research has demonstrated that they can ease comprehension of complex, legal matters for laypeople and legal experts alike. Chapter 4 ends with the inclusion of many examples of visual law, with a focus on contract visualization and design patterns. The latter are generally defined as replicable and systematized solutions to common problems. Particular attention has been devoted to visual patterns that can be employed in legal documents to achieve transparency of information and informed consent, while a related stream of research has focused on design patterns of various nature that support the realization in practice of the abstract principles of privacy by design.

This topic has paved the way to Chapter 5, that revolves around the key role that design can play to effectively apply the principles and values of EU data protection law. Firstly, the principles of data protection by design and by default and their revolutionary inclusion in the GDPR have been described, together with some of their concrete implications. Indeed, design can be used as empowering tool to develop privacy-preserving technologies to achieve privacy-friendly outcomes. Information design and interface design can constitute a subarea of such technologies if considered within the view of compliance with the obligations of transparency and informed consent. We have thus explored how privacy design patterns, especially those borrowed from research on usability and human-computer interaction, can constitute valuable and viable solutions to the many problems that had been identified in Chapter 2.

Therefore, we have mapped the *status quo* of privacy-related communication and consent requests in the online environment (with an analysis limited to webpages) with possible and emerging design patterns that translate GDPR requirements into applicable solutions. Three categories have been identified (i.e. language patterns, visualization patterns, and interaction patterns), whose patterns have been further classified according to one or more of the functional problems identified in Chapter 2 that they aim to solve. Among such patterns, particular relevance for the present project has assumed the data protection icon pattern, that has been explored in depth in the following chapter. Lastly, Chapter 5 has also analyzed the use of design with malicious intent, such as deceiving users and creating privacy-corrosive technologies. We have provided examples of how bad information design can obscure information in unintelligible and hard-to-navigate privacy policies, while deliberate interface design choices (such as default choices) can nudge users towards personal data disclosures or towards consent to certain processing operations.

Chapter 6 has been completely dedicated to DaPIS, the Data Protection Icon Set that constitutes the focus of this research. With the coming into effect of the GDPR, the theoretical discussion and the provision of evidence on how to produce, evaluate, and use icons for data protection have become timely and needed and this chapter aims to constitute a contribution to such debate. Firstly, we have described the extent to which (legal and data protection) icons have idiosyncratic features if compared to other typologies of (legal) visualizations that were introduced in Chapter 4. Although it is commonly believed that such graphical symbols can convey meanings universally, their ease of recognition depends in fact from a combination of factors, that have to be carefully examined when designing and consequently evaluating the icon set. Most prominently, familiarity with a visual representation and familiarity with the underlying referent are fundamental to ensure comprehension of the icon's meaning. However, in the legal sphere, the number of popular symbols is very limited, while it is hard for individuals without legal expertise to be familiar with concepts of data protection. The depiction of concrete objects is also much more effective than the representation of abstract notions. Legibility is also a crucial dimension to determine ease of recognition, while the provision of contextual clues also greatly contribute to support the understanding of the meaning of icons.

We have described previous research where such characteristics were not taken into account in the evaluation phase, thus wrongly suggesting the abandon of those data protection icons that were not readily recognized. In other cases, the evaluation was critical to determine strengths and weaknesses of the icons and to stress the importance of individual characteristics, such as cultural and professional background, in the interpretation process. Considered these previous experiences and the constraints introduced above, the research around DaPIS has resorted to methods and tools borrowed from human-centered design. In Chapter 6 a series of participatory design workshops has been described, in which designers and legal experts, together with computer scientists and interested citizens, collaborated and made use of the reciprocal knowledge and skills to design DaPIS. During these workshops, a profound tension between two different mind sets and goals emerged dramatically: whereas the legal experts deemed fundamental to represent data protection concepts as precisely as possible in order to avoid misinterpretations, the designers fittingly insisted on the simplicity of the icon design for usability reasons. This opposition re-emerged several times also during the evaluation of DaPIS and a mediation between these views has proved challenging.

This chapter has also described the iterative phases of evaluation and consequent vetting of the icon set, that involved users of different demographics and that pointed out the main flows in the icon design. Notwithstanding the precious lessons learned from user research, it is simply inevitable that some graphical symbols will be less transparent than others. Whereas the widest level of recognition should be favored to guide data subjects towards correct interpretation of the icons, a certain threshold of acceptability must be to a certain extent arbitrarily established. Indeed, the chapter ends for a call on standardization of the icon set and education of data subjects to data protection: without these two necessary steps, it is impossible to produce a visual language that will be promptly and flawlessly interpreted across Europe.

Finally, Chapter 7 has provided two sets of contributions. On the one hand, future directions of research have been suggested with the objective of gauging more precisely the effectiveness of DaPIS along dimensions that had not been evaluated in the present research, e.g. effects of training ease, effectiveness of the icon set in context; discriminability across the set elements. On the other hand, the many problems that the development of this project have made emerge were pointed out and elaborated as open questions that deserve further research. In conclusion, the research described in this dissertation has attempted to contribute to the scientific debate on the design and evaluation of data protection icons (i.e. an evidence-based approach as suggested by the WP29) and has set the foundations for future, further investigation into this promising transparency mechanism.

Bibliography

[1] Seminario law & design for privacy.

[2] Seminario law & design for privacy 2.

[3] Une communication juridique impactante avec le legal design. Online at https://www.desmarais-avocats.fr/legal-design-rgpd/.

[4] Evolution of a prototype financial privacy notice. Technical report, Kleimann Communication Group, Inc., 2006.

[5] Functional requirements for bibliographic records. Technical report, International Federation of Library Associations and Institutions, 2009.

[6] Web ontology language (owl). Online at https://www.w3.org/OWL/, December 2012.

[7] Gpen sweep 2017 'user controls over personal information'. Technical report, UK Information Commissioner's Office, October 2017.

[8] Clearly opaque. privacy risks of the internet of things. Technical report, The Internet of Things Privacy Forum, May 2018.

[9] Deceived by design. how tech companies use dark patterns to discourage us from exercising our rights to privacy. Technical report, Forbruker Radet, June 27, 2018.

[10] Digcomp into action. get inspired make it happen. a user guide to the european digital competence framework. Technical report, Joint Research Centre, May 2018.

[11] The juro privacy policy. Online at https://juro.com/policy.html, 2018.

[12] Digcomp: Digital competence framework for citizens, Last update: 20/05/2018.

[13] Martin Abrams. The origins of personal data and its implications for governance. 2014.

[14] Chadia Abras, Diane Maloney-Krichmar, and Jenny Preece. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, 37(4):445–456, 2004.

[15] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, et al. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys (CSUR)*, 50(3):44, 2017.

[16] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Privacy and human behavior in the age of information. *Science*, 347(6221):509–514, 2015.

[17] Alessandro Acquisti and Jens Grossklags. What can behavioral economics teach us about privacy. *Digital Privacy: Theory, Technologies and Practices*, 18:363–377, 2007.

[18] Idris Adjerid, Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Sleights of privacy: Framing, disclosures, and the limits of transparency. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, page 9. ACM, 2013.

[19] William Albert and Thomas Tullis. *Measuring the user experience: collecting, analyzing, and presenting usability metrics.* Newnes, 2013.

[20] Alberto Alemanno and Anne-Lise Sibony. *Nudge and the law: A European perspective.* Bloomsbury Publishing, 2015.

[21] Christopher Alexander. *A pattern language: towns, buildings, construction.* Oxford university press, 1977.

[22] Article 29 Data Protection Working Party. Opinion 10/2004 on more harmonised information provision. Online, November 2004.

[23] Article 29 Data Protection Working Party. Opinion 02/2013 on apps on smart devices 00461/13/en wp 202, February 2013.

[24] Article 29 Data Protection Working Party. Opinion 05/2014 on anonymisation techniques, April 2014.

[25] Article 29 Data Protection Working Party. Opinion 8/2014 on the recent developments of the internet of things. Online, September 2014.

[26] Article 29 Data Protection Working Party. Guidelines on consent under regulation 2016/679, 17/EN WP259, November 2017.

[27] Article 29 Data Protection Working Party. Guidelines on personal data breach notification under regulation 2016/679, October 2017.

[28] Article 29 Data Protection Working Party. Guidelines on transparency under regulation 2016/679, 17/EN WP260, December 2017.

[29] Article 29 Data Protection Working Party. Guidelines on automated individual decision-making and profiling for the purposes of regulation 2016/679 17/en wp251rev.01, February 2018.

[30] Article 29 Data Protection Working Party. Guidelines on transparency under regulation 2016/679, 17/EN WP260 rev.01, April 2018.

[31] Article 29 DPWP. Opinion 15/2011 on the definition of consent wp187. Online, 2011.

[32] Michele M Asprey. *Plain language for lawyers*. Federation Press, 2003.

[33] Canadian Bar Association. Reading the legal world: Literacy and justice in canada. report of the canadian bar association task force on legal literacy. Technical report, Ottawa, 1992.

[34] Tara Athan, Harold Boley, Guido Governatori, Monica Palmirani, Adrian Paschke, and Adam Wyner. Oasis legalruleml. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, pages 3–12. ACM, 2013.

[35] Tara Athan, Guido Governatori, Monica Palmirani, Adrian Paschke, and Adam Wyner. LegalRuleML: Design principles and foundations. In *Reasoning Web International Summer School*, pages 151–188. Springer International Publishing, 2015.

[36] Gioele Barabucci, Luca Cervone, Angelo Di Iorio, Monica Palmirani, Silvio Peroni, and Fabio Vitali. Managing semantics in xml vocabularies: an experience in the legal and legislative domain. In *Proceedings of Balisage: The markup conference*, volume 5, 2010.

[37] Gioele Barabucci, Luca Cervone, Monica Palmirani, Silvio Peroni, and Fabio Vitali. Multi-layer markup and ontological structures in Akoma Ntoso. In *AI Approaches to the Complexity of Legal Systems. Complex Systems, the Semantic Web, Ontologies, Argumentation, and Dialogue*, pages 133–149. Springer, 2010.

[38] Gioele Barabucci, Angelo Di Iorio, Francesco Poggi, and Fabio Vitali. Integration of legal datasets: from meta-model to implementation. In *Proceedings of International Conference on Information Integration and Web-based Applications & Services*, page 585. ACM, 2013.

[39] Thomas D Barton, Gerlinde Berger-Walliser, and Helena Haapio. Visualization: seeing contracts for what they are, and what they could become. *JL Bus. & Ethics*, 19:47, 2013.

[40] Omri Ben-Shahar and Carl E Schneider. The failure of mandated disclosure. *University of Pennsylvania Law Review*, pages 647–749, 2011.

[41] Omri Ben-Shahar and Carl E Schneider. *More than you wanted to know: The Failure of Mandated Disclosure*. Princeton University Press, 2014.

[42] Gerlinde Berger-Walliser, Thomas D Barton, and Helena Haapio. From visualization to legal design: A collaborative and creative process. *American Business Law Journal*, 54(2):pp. 347–392, 2017.

[43] Gerlinde Berger-Walliser, Robert C Bird, and Helena Haapio. Promoting business success through contract visualization. *JL Bus. & Ethics*, 17:55, 2011.

[44] Jaspreet Bhatia, Travis D Breaux, Joel R Reidenberg, and Thomas B Norton. A theory of vagueness and privacy risk perception. In *Requirements Engineering Conference (RE), 2016 IEEE 24th International*, pages 26–35. IEEE, 2016.

[45] Maria Angela Biasiotti. Semantic resources for managing legislative information. In *Legislative XML for the Semantic Web*, pages 151–172. Springer, 2011.

[46] Volker Boehme-Nessler. *Pictorial law: modern law and the power of pictures*. Springer Science & Business Media, 2010.

[47] Danielle Bond. Australia's first visual employment contracts launched. Online at `https://www.aurecongroup.com/about/latest-news/2018/may/visual-employment-contract`, May 5 2018.

[48] Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp, and Stefan Pfattheicher. Tales from the dark side: Privacy dark strategies and privacy dark patterns. *Proceedings on Privacy Enhancing Technologies*, 2016(4):237–254, 2016.

[49] Marietjie Botes. Using comics to communicate legal contract cancellation. *The Comics Grid: Journal of Comics Scholarship*, 7, 2017.

[50] J. A. P. J. Breuker, R. Hoekstra, K. van den Berg, R. Rubino, Giovanni Sartor, Monica Palmirani, Adam Wyner, and T. Bench-Capon. Owl ontology of basic legal concepts (lkif-core). estrella: Deliverable. Technical report, Amsterdam, UVA, 2007.

[51] Tim Brown. Change by design. 2009.

[52] Colette Brunschwig. *Visualisierung von Rechtsnormen: legal design*. PhD thesis, University of Zürich, 2001.

[53] Jenna Burrell. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1):2053951715622512, 2016.

[54] Christoph Busch. *Research handbook on EU consumer and contract law*, chapter The Future of Pre-Contractual Information Duties: From Behavioural Insights to Big Data. Edward Elgar Publisher, 2016.

[55] Christoph Busch. Implementing personalized law: Personalized disclosures in consumer law and privacy law. *University of Chicago Law Review*, 2018.

[56] Peter Butt. Legalese versus plain language. *Amicus Curiae*, 2001(35):28–32, 2012.

[57] Patrice Caire, Nicolas Genon, Patrick Heymans, and Daniel L Moody. Visual notation design 2.0: Towards user comprehensible requirements

engineering notations. In *Requirements Engineering Conference (RE), 2013 21st IEEE International*, pages 115–124. IEEE, 2013.

[58] Julio C Caiza, Yod-Samuel Martín, Jose M Del Alamo, and Danny S Guamán. Organizing design patterns for privacy: a taxonomy of types of relationships. In *Proceedings of the 22nd European Conference on Pattern Languages of Programs*, page 32. ACM, 2017.

[59] M Ryan Calo. Against notice skepticism in privacy (and elsewhere). *notre dame law review*, 87:3, 2012.

[60] Tom Calver and Joe Miller. Social site terms tougher than dickens, July 6 2018.

[61] Eoin Carolan. The continuing problems with online consent under the eu's emerging data protection principles. *Computer Law & Security Review*, 32(3):462–473, 2016.

[62] Ann Cavoukian. Privacy by design. the 7 foundational principles. *Take the challenge. Information and privacy commissioner of Ontario, Canada*, 2009.

[63] Ann Cavoukian. Privacy by design in law, policy and practice. *A white paper for regulators, decision-makers and policy-makers*, 2011.

[64] Eun Kyoung Choe, Jaeyeon Jung, Bongshin Lee, and Kristie Fisher. Nudging people away from privacy-invasive mobile apps through visual framing. In *IFIP Conference on Human-Computer Interaction*, pages 74–91. Springer, 2013.

[65] Andrea E Cladis. A shifting paradigm: An evaluation of the pervasive effects of digital technologies on language expression, creativity, critical thinking, political discourse, and interactive processes of human communications. *E-Learning and Digital Media*, page 2042753017752583.

[66] James Clark et al. Xsl transformations (xslt). *World Wide Web Consortium (W3C). URL http://www. w3. org/TR/xslt*, page 103, 1999.

[67] CNIL. The cnil's restricted committee imposes a financial penalty of 50 million euros against google llc. Online at https://www.cnil.fr/en/
cnils-restricted-committee-imposes-financial-penalty-50-million-eu
mkt_tok=eyJpIjoiTkRjNVpUUmhOakkwTkRnMiIsInQiOiJsc1h3Y3VmZTVIN3RlVm

[68] CNN. Senator to zuckerberg: Your user agreement sucks. Online at https://edition.cnn.com/videos/politics/2018/04/10/john-kennedy-zuckerberg-user-agreement-sucks-erin-sot.cnn, April 10 2018.

[69] Michael Colesky, Jaap-Henk Hoepman, and Christiaan Hillen. A critical analysis of privacy design strategies. In *Security and Privacy Workshops (SPW), 2016 IEEE*, pages 33–40. IEEE, 2016.

[70] UK Children's Commissioner. Growing up digital. a report from the children commissioner's growing up digital taskforce. Technical report, January 2017.

[71] Giuseppe Contissa, Koen Docter, Francesca Lagioia, Marco Lippi, Hans-W Micklitz, Przemyslaw Palka, Giovanni Sartor, and Paolo Torroni. Claudette meets gdpr. automating the evaluation of privacy policies using artificial intelligence. Technical report, BEUC, 2018.

[72] Lorrie Faith Cranor, Praveen Guduru, and Manjula Arjula. User interfaces for privacy agents. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 13(2):135–178, 2006.

[73] John W Creswell, Vicki L Plano Clark, Michelle L Gutmann, and William E Hanson. Advanced mixed methods research designs. *Handbook of mixed methods in social and behavioral research*, 209:240, 2003.

[74] Nathan Crilly, David Good, Derek Matravers, and P John Clarkson. Design as communication: exploring the validity and utility of relating intention to interpretation. *Design Studies*, 29(5):425–457, 2008.

[75] George Danezis, Josep Domingo-Ferrer, Marit Hansen, Jaap-Henk Hoepman, Daniel Le Metayer, Rodica Tirtea, and Stefan Schiffner. Privacy and data protection by design-from policy to engineering. 2015.

[76] David B Daniel and William Douglas Woody. E-textbooks at what cost? performance and use of electronic v. print texts. *Computers & Education*, 62:18–23, 2013.

[77] Janet Davis and Lisa Nathan. *Handbook of Ethics, Values and Technological Design*, chapter Value Sensitive Design: Applications, Adaptations and Critiques, pages 11–40. Dordrecht Springer, 2015.

[78] Conseil de l'Europe. *Convention for the protection of individuals with regard to automatic processing of personal data*, volume 108. Council of Europe, 1981.

[79] Robert de Rooy. Comic contract - clemengold fruit picker agreement. Online at `https://www.yumpu.com/en/document/view/55339897/indigo-letsitele-comic-contract-booklet-draft-smlr/2`, 2016.

[80] Clarisse Sieckenius de Souza. Semiotic engineering: bringing designers and users together at interaction time. *Interacting with Computers*, 17(3):317–341, 2005.

[81] Clarisse Sieckenius de Souza, Carla Faria Leitão, Raquel Oliveira Prates, Sílvia Amélia Bim, and Elton José da Silva. Can inspection methods generate valid new knowledge in hci? the case of semiotic inspection. *International Journal of Human-Computer Studies*, 68(1-2):22–40, 2010.

[82] Sebastian Deterding. Gamification: designing for motivation. *interactions*, 19(4):14–17, 2012.

[83] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pages 9–15. ACM, 2011.

[84] Robert Dewar. *Visual information for everyday use: Design and research perspectives*, chapter Design and evaluation of public information symbols, pages 285–303. Taylor and Francis London, 1999.

[85] Michal Dudek. Why are words not enough? or a few remarks on traffic signs. In *Problems of Normativity, Rules and Rule-Following*, pages 363–372. Springer, 2015.

[86] Easyjet. Privacy policy and our privacy promise. Online at `https://www.easyjet.com/en/policy/privacy-promise`. Last accessed: July 1, 2018.

[87] Dag Elgesem. Privacy, respect for persons, and risk. *Philosophical Perspectives on Computer-Mediated Communication. State University of New York Press, Albany*, pages 45–66, 1996.

[88] Timothy Endicott. Vagueness in law. 2000.

[89] Timothy Endicott. Law is necessarily vague. *Legal theory*, 7(4):379–385, 2001.

[90] Martin J Eppler and Remo A Burkhard. Knowledge visualization. Technical report, Università della Svizzera italiana, 2004.

[91] Samson Esayas, Tobias Mahler, and Kevin McGillivray. Is a picture worth a thousand terms? visualising contract terms and data protection requirements for cloud computing users. In *International Conference on Web Engineering*, pages 39–56. Springer, 2016.

[92] EU High Level Group of Experts on Literacy. Final report, september 2012. Technical report, European Commission, 2012.

[93] European Commission. Communication to the european parliament, the council, the economic and social committee and the committee of the regions: a comprehensive approach on personal data protection in the european union. com(2010) 609 final, November, 4 2010.

[94] European Convention. Charter of fundamental rights of the european union (2012/c 326/02).

[95] European Data Protection Supervisor. Opinion of the european data protection supervisor on the communication from the commission to the european parliament, the council, the economic and social committee and the committee of the regions - ”a comprehensive approach on personal data protection in the european union”, January, 14 2011.

[96] European Data Protection Supervisor. Preliminary opinion of the european data protection supervisor privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the digital economy, March 2014.

[97] European Data Protection Supervisor. Opinion 4/2015 towards a new digital ethics. Online, September 2015.

[98] European Data Protection Supervisor. Opinion 05/2018. preliminary opinion on privayc by design, May 2018.

[99] European Parliament and Council of European Union. Directive (EU) 95/46/ec of the european parliament and of the council on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official Journal L 281 , 23/11/1995 P. 0031 - 0050, October 1995.

[100] European Parliament and Council of European Union. Directive 2002/58/ec of the european parliament and of the council of 12 july

2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (directive on privacy and electronic communications). O.J. L 201, 31.7.2002, p. 37–47, 2002.

[101] European Parliament and Council of European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). O.J. L 119, 4.5.2016, p. 1–88, 2016.

[102] European Parliament. Committee on Civil Liberties, Justice and Home Affairs. Draft report on the proposal for a regulation of the european parliament and of the council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (general data protection regulation) [com(2012)0011 - c7-0025/2012 - 2012/0011 (cod)], 21 November 2013.

[103] Benjamin Fabian, Tatiana Ermakova, and Tino Lentz. Large-scale readability analysis of privacy policies. In *Proceedings of the International Conference on Web Intelligence*, pages 18–25. ACM, 2017.

[104] FEDMA. Code of practice for the use of personal data, 1998.

[105] Neal Feigenson and Christina Spiesel. *Law on display: The digital transformation of legal persuasion and judgment*. NYU Press, 2009.

[106] S. Fischer-H́ubner, E. Ẃastlund, and H. Zwingelberg. Ui prototypes: Policy administration and presentation version 1, deliverable d4.3.1 of the ec fp7 project primelife. Technical report, PrimeLife Consortium, 2009.

[107] S. Fischer-H́ubner, E. Ẃastlund, and H. Zwingelberg. Ui prototypes: Policy administration and presentation version 2, deliverable d4.3.2 of

the ec fp7 project primelife. Technical report, PrimeLife Consortium, 2010.

[108] Simone Fischer-Hübner, C Köffel, JS Pettersson, P Wolkerstorfer, C Graf, LE Holtz, U König, H Hedbom, and B Kellermann. Hci pattern collection–version 2. *Priv. Identity Manag. Eur. Life*, 61, 2010.

[109] Uwe Flick. *An introduction to qualitative research*. Sage Publications Limited, 2018.

[110] Center for Urban Pedagogy. I got arrested! now what? a guide to the juvenile justice system, 2010.

[111] The Center for Urban Pedagogy. Vendor power: a guide to street vending in new york city. Online at `http://welcometocup.org/Projects/MakingPolicyPublic/VendorPower`, 2009.

[112] Pirjo-Leena Forsström, Helena Haapio, and Stefania Passera. Fair design jam: A case study on co-creating communication about fair data principles. In Erich Schweighofer et al. (Eds.), editor, *Trend and Communities of Legal Informatics. Proceedings of the 20th International Legal Informatics Symposium IRIS 2017*, pages pp. 433–440, Wien, 2017.

[113] Charles Fried and David Ferdinand Schoeman. Philosophical dimensions of privacy: An anthology, 1984.

[114] Batya Friedman and Peter H Kahn Jr. Human values, ethics, and design. In Julie A. Jacko Andrew Sears, editor, *The human-computer interaction handbook*, pages 1223–1248. CRC Press, 2007.

[115] Batya Friedman, Peyina Lin, and Jessica K Miller. Informed consent by design. *Security and Usability*, (2001):503–530, 2005.

[116] A Michael Froomkin. The death of privacy. *Stan. L. Rev.*, 52:1461, 1999.

[117] Erich Gamma. *Design patterns: elements of reusable object-oriented software*. Pearson Education India, 1995.

[118] Aldo Gangemi, Silvio Peroni, David Shotton, and Fabio Vitali. The publishing workflow ontology (pwo). *Semantic Web*, 8(5):703–718, 2017.

[119] Aldo Gangemi and Valentina Presutti. Ontology design patterns. In *Handbook on ontologies*, pages 221–243. Springer, 2009.

[120] Anton Geist, Colette Brunschwig, Friedrich Lachmayer, and Günther Schefbeck. Multisensory law and legal informatics: a comparison of how these legal disciplines relate to visual law. In *Strukturierung der Juristischen Semantik - Structuring Legal Semantics*. Editions Weblaw, 2011.

[121] Marta Ghio, Matilde Maria Serena Vaghi, and Marco Tettamanti. Fine-grained semantic categorization across the abstract and concrete domains. *PloS one*, 8(6):e67090, 2013.

[122] Joshua Gomez, Travis Pinnick, and Ashkan Soltani. Privacy coding methodology. Online, 2009.

[123] Gloria González Fuster. How uninformed is the average data subject? a quest for benchmarks in eu personal data protection. *IDP. Revista de Internet, Derecho y Política*, (19), 2014.

[124] Ravindra S Goonetilleke, Heloisa Martins Shih, and JULIEN FRITSCH. Effects of training and representational characteristics in icon design. *International Journal of Human-Computer Studies*, 55(5):741–760, 2001.

[125] Cornelia Graf, Christina Hochleitner, Peter Wolkerstorfer, Julio Angulo, Simone Fischer-Hübner, and Erik Wästlund. Final hci research project. Technical report, PrimeLife Consortium, 2011.

[126] Thomas R Gruber. A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2):199–220, 1993.

[127] The Guardian. The cambridge analytica files. a year-long investigation into facebook, data, and influencing elections in the digital age. Online at https://www.theguardian.com/news/series/cambridge-analytica-files, 2018.

[128] The Guardian. Privacy policy. Online at https://www.theguardian.com/help/privacy-policy. Last accessed: 22 February 2018, May 24 2018.

[129] Nicola Guarino and Christopher Welty. Evaluating ontological decisions with ontoclean. *Communications of the ACM*, 45(2):61–65, 2002.

[130] Helena Haapio. *Next generation contracts: a paradigm shift.* Lexpert, 2013.

[131] Helena Haapio. Bringing design thinking to contract design. Online, January 2014.

[132] Helena Haapio, G Berger-Walliser, B Walliser, and Katri Rekola. Time for a visual turn in contracting. *Journal of Contract Management*, 10:49–58, 2012.

[133] Helena Haapio and Margaret Hagan. Design patterns for contracts. In *Networks. Proceedings of the 19th International Legal Informatics Symposium IRIS*, 2016.

[134] Helena Haapio, Margaret Hagan, Monica Palmirani, and Arianna Rossi. Legal design patterns for privacy. In Eric Schweighofer and et al., editors, *Data Protection / LegalTech Proceedings of the 21st International Legal Informatics Symposium IRIS 2018*, pages pp. 445–450. Editions Weblaw, 2018.

[135] Helena Haapio and Stefania Passera. Visual law: what lawyers need to learn from information designers. *VoxPopuLII, Legal Information Institute*, 2013.

[136] Helena Haapio and Stefania Passera. Contracts as interfaces: Exploring visual representation patterns in contract design. In R.A. Dolin M. J. Katz and M. Bommarito, editors, *Legal Informatics*. UK: Cambridge University Press, Forthcoming.

[137] Helena Haapio, Daniela Alina Plewe, and Robert DeRooy. Next generation deal design: Comics and visual platforms for contracting. In *Networks. Proceedings of the 19th International Legal Informatics Symposium IRIS*, pages 373–380, 2016.

[138] Helena Haapio, Daniela Alina Plewe, and Robert deRooy. Contract continuum: From text to images, comics, and code. In Erich Schweighofer et al. (Eds.), editor, *Trend and Communities of Legal Informatics. Proceedings of the 20th International Legal Informatics Symposium IRIS 2017*, pages pp. 411–418, 2017.

[139] Philipp Hacker. Personalizing eu private law. from disclosures to nudges and mandates. *25 European review of Private Law*, 2017.

[140] Margaret Hagan. Legal design lab: a new generation of legal services.

[141] Margaret Hagan. Law by design. http://www.lawbydesign.co, 2017.

[142] Margaret Hagan. Rethinking data privacy communication design: 3 big questions from bologna. Online: `https://medium.com/legal-design-and-innovation/rethinking-data-privacy-communication-design-3-big-questions-from-` April 29 2018.

[143] Margaret Hagan and Helena Haapio. Contract design pattern library. Online at `http://www.`

legaltechdesign.com/communication-design/legal-design-pattern-libraries/contracts/, 2016.

[144] Marit Hansen. Putting privacy pictograms into practice-a european perspective. *GI Jahrestagung*, 154, 2009.

[145] Hamza Harkous, Kassem Fawaz, Rémi Lebret, Florian Schaub, Kang G Shin, and Karl Aberer. Polisis: Automated analysis and presentation of privacy policies using deep learning. *arXiv preprint arXiv:1802.02561*, 2018.

[146] Hamza Harkous, Kassem Fawaz, Kang G Shin, and Karl Aberer. Pribots: Conversational privacy with chatbots. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. USENIX Association, 2016.

[147] Woodrow Hartzog. Website design as contract. *American University Law Review*, 60:1635, 2011.

[148] Woodrow Hartzog. *Privacy's Blueprint: The Battle to Control the Design of New Technologies*. Harvard University Press, 2018.

[149] Genevieve Helleringer and Anne-Lise Sibony. European consumer protection through the behavioral lens. *Colum. J. Eur. L.*, 23:607, 2016.

[150] Jon Hicks. *The icon handbook*. Five Simple Steps, 2011.

[151] Robert A Hillman and Jeffrey J Rachlinski. Standard-form contracting in the electronic age. *NYUL Rev.*, 77:429, 2002.

[152] Mike Hintze. In defense of the long privacy statement. *Maryland Law Review*, 76(4):1044–1085, 2017.

[153] Mark Hochhauser. Why patients won't understand their hipaa privacy notices. Online at https://www.privacyrights.org/blog/why-patients-wont-understand-their-hipaa-privacy-notices-hochhau 2003.

[154] Jaap-Henk Hoepman. Privacy design strategies. In *IFIP International Information Security Conference*, pages 446–459. Springer, 2014.

[155] Leif-Erik Holtz, Katharina Nocun, and Marit Hansen. Towards displaying privacy information with icons. In Camenisch Jan, Crispo Bruno, Fischer-Hübner Simone, Leenes Ronald, and Russello Giovanni, editors, *IFIP PrimeLife International Summer School on Privacy and Identity Management for Life*, pages 338–348. Springer, 2010.

[156] William K Horton. *The icon book: Visual symbols for computer systems and documentation*. John Wiley & Sons, Inc., 1994.

[157] Jinghui Hou, Justin Rashid, and Kwan Min Lee. Cognitive map or medium materiality? reading on paper and screen. *Computers in Human Behavior*, 67:84–94, 2017.

[158] Renato Iannella. Open digital rights language (odrl) version 1.1. Online at `https://www.w3.org/TR/odrl/`, September 2002.

[159] Renato Iannella and Adam Finden. Privacy awareness: Icons and expression for social networks. In *Proceedings of the 8th Virtual Goods Workshop and the 6th ODRL Workshop*, pages 1–15, 2010.

[160] IAPP and Create with Context. The ux guide to getting consent. Online at `https://iapp.org/store/books/a191a000002FUZKAA4/`, 2017.

[161] ICO. Deleting personal data. Technical report, 2014.

[162] ICO. Encryption, 2016.

[163] ETSI (European Telecommunications Standards Institute). Human factors (hf); framework for the development, evaluation and selection of graphical symbols. eg 201 379 v1.1.1 (1998-12). Online at: http://www.etsi.org/, 1998.

[164] Sarah Isherwood. Graphics and semantics: The relationship between what is seen and what is meant in icon design. In *International Conference on Engineering Psychology and Cognitive Ergonomics*, pages 197–205. Springer, 2009.

[165] Sarah J Isherwood, Siné JP McDougall, and Martin B Curry. Icon identification in context: The changing role of icon characteristics with user experience. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(3):465–476, 2007.

[166] ISO. Iso 9241-210:2010. ergonomics of human-system interaction – part 210: Human-centred design for interactive systems, 2010.

[167] ISO. Iso 9186-1:2014. graphical symbols – test methods – part 1: Method for testing comprehensibility, 2014.

[168] ISO. Iso 9186-3:2014. graphical symbols — test methods part 3: Method for testing symbol referent association, 2014.

[169] ISO. Iso 31000:2018 – risk management, 2018.

[170] Roman Jakobson. Closing statement: Linguistics and poetics. *Style in language*, pages pp. 350–377, 1960.

[171] Milena Janic, Jan Pieter Wijbenga, and Thijs Veugen. Transparency enhancing tools (tets): an overview. In *2013 Third Workshop on Socio-Technical Aspects in Security and Trust*, pages 18–25. IEEE, 2013.

[172] Carlos Jensen and Colin Potts. Privacy policies as decision-making tools: an evaluation of online privacy notices. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 471–478. ACM, 2004.

[173] Matthew Kay. Techniques and heuristics for improving the visual design of software agreements. Master's thesis, University of Waterloo, 2010.

[174] Matthew Kay and Michael Terry. Textured agreements: re-envisioning electronic consent. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, page 13. ACM, 2010.

[175] Adrian Keating and Camilla Baasch Andersen. A graphic contract: Taking visualisation in contracting a step further. *Journal of Strategic Contracting and Negotiation*, 2(1-2):10–18, 2016.

[176] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W Reeder. A nutrition label for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, page 4. ACM, 2009.

[177] Patrick Gage Kelley, Lucian Cesca, Joanna Bresee, and Lorrie Faith Cranor. Standardizing privacy notices: an online study of the nutrition label approach. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, pages 1573–1582. ACM, 2010.

[178] Bart P Knijnenburg and David Cherry. Comics as a medium for privacy notices. In *WSF@ SOUPS*, 2016.

[179] Klaus Krippendorff. *The semantic turn: A new foundation for design.* crc Press, 2005.

[180] Klaus Krippendorff and Reinhart Butter. Product semantics-exploring the symbolic qualities of form. *Departmental Papers (ASC)*, page 40, 1984.

[181] Jean-Baptiste Lamy and Lina F Soualmia. Formalization of the semantics of iconic languages: An ontology-based method and four semantic-powered applications. *Knowledge-Based Systems*, 135:159–179, 2017.

[182] Marc Langheinrich. A privacy awareness system for ubiquitous computing environments. In *international conference on Ubiquitous Computing*, pages 237–245. Springer, 2002.

[183] Legal Design Lab. Design workshop for eu general data protection regulation. http://www.legaltechdesign.com/design-workshop-for-eu-general-data-protection-regulation/, July 2017.

[184] Lawrence Lessig. *Code: And other laws of cyberspace.* ReadHowYouWant. com, 2009.

[185] Commission Nationale Informatique & libertés. Connected vehicles and personal data, October 2017.

[186] Linkedin. Privacy policy. Online at `https://www.linkedin.com/legal/privacy-policy`. Last accessed: July 3, 2018, May 8 2018.

[187] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Aerin Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. Follow my recommendations: A personalized assistant for mobile app permissions. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, 2016.

[188] Ewa Luger, Stuart Moran, and Tom Rodden. Consent for all: revealing the hidden complexity of terms and conditions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 2687–2696. ACM, 2013.

[189] Francisco Lupiáñez-Villanueva, George Gaskell, Pietro Tornese, José Vila, Yolanda Gómez, Anthony Allen, Giuseppe Veltri, and Cristiano Codagnone. Behavioural study on the transparency of online platforms final report. Technical report, European Commission, April 2018.

[190] Wainer Lusoli, Margherita Bacigalupo, Francisco Lupiáñez-Villanueva, Norberto Nuno Gomes de Andrade, Shara Monteleone, and Ioannis Maghiros. Pan-european survey of practices, attitudes and policy preferences as regards personal identity data management. 2012.

[191] John Lyons. Semantics vol. ii. *Cambridge CUP*, 1977.

[192] Richard Mabey. Privacy by design: Building a privacy policy people actually want to read. Online at https://www.artificiallawyer.com/2018/05/02/privacy-by-design-building-a-privacy-policy-people-actually-want-t May 2 2018.

[193] Mary Madden, Amanda Lenhart, Sandra Cortesi, Urs Gasser, Maeve Duggan, Aaron Smith, and Meredith Beaton. Teens, social media, and privacy. *Pew Research Center*, 21:2–86, 2013.

[194] Connie Malamed. *Visual language for designers: principles for creating graphics that people understand*. Rockport Publishers, 2009.

[195] Anne Mangen, Bente R Walgermo, and Kolbjørn Brønnick. Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International journal of educational research*, 58:61–68, 2013.

[196] Alessandro Mantelero. The future of consumer data protection in the eu re-thinking the “notice and consent” paradigm in the new era of predictive analytics. *Computer Law & Security Review*, 30(6):643–660, 2014.

[197] Helia Marreiros, Richard Gomer, Michael Vlassopoulos, Mirco Tonin, et al. Exploring user perceptions of online privacy disclosures. *In proceedings IADIS International Conference WWW/Internet - ICWI 2015, 07 - 08 Jun 2015*, 2015.

[198] Aleecia M McDonald and Lorrie Faith Cranor. The cost of reading privacy policies. *ISJLP*, 4:543, 2008.

[199] M. Mehldau. Iconset for data-privacy declarations v 0.1, 2007.

[200] David Mellinkoff. *The language of the law*. Wipf and Stock Publishers, 2004.

[201] Naomi Mezey. Image cannot speak for itself: Film, summary judgment, and visual literacy, the. *Val. UL Rev.*, 48:1, 2013.

[202] George R Milne and Mary J Culnan. Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices. *Journal of Interactive Marketing*, 18(3):15–29, 2004.

[203] Christopher F Mondschein. Some iconoclastic thoughts on the effectiveness of simplified notices and icons for informing individuals as proposed in article 12(1) and (7) gdpr. *European Data Protection Law Review*, 2(4):507–520, 2016.

[204] Shara Monteleone. Addressing the "failure" of informed consent in online data protection: Learning the lessons from behaviour-aware regulation. *Syracuse Journal of International Law and Commerce*, 43:69–119, 2015.

[205] Shara Monteleone, René van Bavel, Nuria Rodríguez-Priego, and Gabriele Esposito. Nudges to privacy behaviour: Exploring an alternative approach to privacy notices. *JRC Science and Policy Report*, 2015.

[206] Daniel Moody. The "physics" of notations: toward a scientific basis for constructing visual notations in software engineering. *IEEE Transactions on Software Engineering*, 35(6):756–779, 2009.

[207] Timothy Morey, Theodore Forbath, and Allison Schoop. Customer data: Designing for transparency and trust. *Harvard Business Review*, 93(5):96–105, 2015.

[208] Ben Moskowitz and Aza Raskin. Privacy icons.

[209] Michael D Murray. Visual rhetoric and visual narrativity in five sections of a brief. Available at SSRN: https://ssrn.com/abstract=2460357 or http://dx.doi.org/10.2139/ssrn.2460357, 2014.

[210] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Cranor, and Norman Sadeh. Privacy expectations and preferences in an iot world. In *Symposium on Usable Privacy and Security (SOUPS)*, 2017.

[211] Helen Nissenbaum. A contextual approach to privacy online. *Daedalus*, 140(4):32–48, 2011.

[212] Donald A Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[213] Charles B Nutting. Graphic law. *American Bar Association Journal*, 50:pp. 780–781, 1964.

[214] Jonathan A Obar and Anne Oeldorf-Hirsch. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. 2016.

[215] Andreas Oehler and Stefan Wendt. Good consumer information: The information paradigm at its (dead) end? *Journal of Consumer Policy*, 40(2):179–191, 2017.

[216] Council of the European Union. Council directive 93/13/eec of 5 april 1993 on unfair terms in consumer contracts. OJ L 95, 21.4.1993, p. 29–34, 1993.

[217] Alessandro Oltramari, Dhivya Piraviperumal, Florian Schaub, Shomir Wilson, Norman Sadeh, and Joel Reidenberg. Privonto: a semantic framework for the analysis of privacy policies. *Semantic Web Journal*, 2016.

[218] Allan Paivio. *Mental representations: A dual coding approach*. Oxford University Press, 1990.

[219] Monica Palmirani and Luca Cervone. Measuring the complexity of the legal order over time. In *AI Approaches to the Complexity of Legal Systems*, pages 82–99. Springer, 2014.

[220] Monica Palmirani, Michele Martoni, Arianna Rossi, Cesare Bartolini, and Livio Robaldo. Legal ontology for modelling gdpr concepts and norms. In *Proceedings of JURIX International Conference on Legal Knowledge and Information Systems*, 2018.

[221] Monica Palmirani, Michele Martoni, Arianna Rossi, Cesare Bartolini, and Livio Robaldo. PrOnto: Privacy Ontology for Legal Compliance. In *Accepted at the 18th European Conference on Digital Government*, 2018.

[222] Monica Palmirani, Michele Martoni, Arianna Rossi, Cesare Bartolini, and Livio Robaldo. PrOnto: Privacy Ontology for Legal Reasoning. In *Accepted at eGovis 2018, the 7th International Conference on Electronic Government and the Information Systems Perspective*, 2018.

[223] Monica Palmirani, Arianna Rossi, Michele Martoni, and Margaret Hagan. A methodological framework to design a machine-readable privacy icon set. In Eric Schweighofer and et al., editors, *Data Protection / LegalTech Proceedings of the 21st International Legal Informatics Symposium IRIS 2018*, pages pp. 451 – 454, Wien, 2018. Editions Weblaw.

[224] Monica Palmirani, Roger Sperberg, Grant Vergottini, and Fabio Vitali. Akoma ntoso version 1.0 part 1: Xml vocabulary. oasis committee specification draft 02 / public review draft 02. Technical report, OASIS Committee Specification, May 2016.

[225] Monica Palmirani and Fabio Vitali. Akoma-Ntoso for legal documents. In *Legislative XML for the semantic Web*, pages 75–100. Springer, 2011.

[226] Monica Palmirani and Fabio Vitali. *Legislative XML: Principles and technical tools*. Inter-American Development Bank, 2012.

[227] European Parliament and Council of European Union. Directive2011/83/eu of the european parliament and of the council of 25 october 2011 on consumer rights amending council directive 93/13/eec and directive 1999/44/ec of the european parliament and of the council and repealing council directive 85/577/eec and directive 97/7/ec of the european parliament and of the council. OJ L 304, 22.11.2011, p. 64–88, 2011.

[228] Frank Pasquale. *The black box society: The secret algorithms that control money and information*. Harvard University Press, 2015.

[229] Stefania Passera. Enhancing contract usability and user experience through visualization-an experimental evaluation. In *16th International Conference on Information Visualisation*, pages 376–382. IEEE, 2012.

[230] Stefania Passera. Beyond the wall of text: How information design can make contracts user-friendly. In *International Conference of Design, User Experience, and Usability*, pages 341–352. Springer, 2015.

[231] Stefania Passera. *Beyond the wall of contract text. Visualizing contracts to foster understanding and collaboration within and across organizations*. PhD thesis, Aalto University, 2017.

[232] Stefania Passera. Flowcharts, swimlanes, and timelines. *Journal of Business and Technical Communication*, 2017.

[233] Stefania Passera and Helena Haapio. Transforming contracts from legal rules to user-centered communication tools: a human-information interaction challenge. *Communication Design Quarterly Review*, 1(3):38–45, 2013.

[234] Stefania Passera, Helena Haapio, and Thomas D Barton. Innovating contract practices: merging contract design with information design. In *Proceedings of the 2013 Academic Forum on Integrating Law and Con-*

tract Management: Proactive, Preventive and Strategic Approaches, 2013.

[235] Stefania Passera, Helena Haapio, and Michael Curtotti. Making the meaning of contracts visible–automating contract visualization. In *Proceedings of the 17th International Legal Informatics Symposium IRIS 2014*, 2014.

[236] Stefania Passera, Soile Pohjonen, Katja Koskelainen, and Suvi Anttila. User-friendly contracting tools–a visual guide to facilitate public procurement contracting. In *Proceedings of the IACCM Academic Forum on Contract and Commercial Management*, 2013.

[237] Charles Sanders Peirce. Logic as semiotic: The theory of signs. 1902.

[238] Silvio Peroni, Monica Palmirani, and Fabio Vitali. Undo: The united nations system document ontology. In *International Semantic Web Conference*, pages 175–183. Springer, 2017.

[239] John Sören Pettersson. A brief evaluation of icons in the first reading of the european parliament on com (2012) 0011,. *Privacy and Identity Management for the Future Internet in the Age of Globalisation. Privacy and Identity 2014. IFIP Advances in Information and Communication Technology*, 457, 2014.

[240] Rune Pettersson. Information design–principles and guidelines. *Journal of Visual Literacy*, 29(2):167–182, 2010.

[241] Travis Pinnick. Privacy short notice design, 2011.

[242] Irene Pollach. A typology of communicative strategies in online privacy policies: Ethics, power and informed consent. *Journal of Business Ethics*, 62(3):221, 2005.

[243] Maria Popova. Mozilla's privacy icons: A visual language for privacy data rights, 2011.

[244] Alexandre Porion, Xavier Aparicio, Olga Megalakaki, Alisson Robert, and Thierry Baccino. The impact of paper-based versus computerized presentation on text comprehension and memorization. *Computers in Human Behavior*, 54:569–576, 2016.

[245] Elizabeth G Porter. Taking images seriously. *Columbia Law Review*, 114(7):1687–1782, 2014.

[246] Robert W Proctor, M Athar Ali, and Kim-Phuong L Vu. Examining usability of web privacy policies. *Intl. Journal of Human–Computer Interaction*, 24(3):307–328, 2008.

[247] Robert W Reeder, Patrick Gage Kelley, Aleecia M McDonald, and Lorrie Faith Cranor. A user study of the expandable grid applied to p3p privacy policy visualization. In *Proceedings of the 7th ACM workshop on Privacy in the electronic society*, pages 45–54. ACM, 2008.

[248] Joel R Reidenberg, Jaspreet Bhatia, Travis D Breaux, and Thomas B Norton. Ambiguity in privacy policies and the impact of regulation. *The Journal of Legal Studies*, 45(S2):S163–S190, 2016.

[249] Joel R Reidenberg, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T Graves, Fei Liu, Aleecia McDonald, Thomas B Norton, and Rohan Ramanath. Disagreeable privacy policies: Mismatches between meaning and users' understanding. *Berkeley Tech. LJ*, 30:39, 2015.

[250] Anna Richter and Mary L Courage. Comparing electronic and paper storybooks for preschoolers: attention, engagement, and recall. *Journal of Applied Developmental Psychology*, 48:92–102, 2017.

[251] Neil Robinson, Hans Graux, Maarten Botterman, and Lorenzo Valeri. Review of the european data protection directive (sponsored by the information commissioner's officer). Technical report, RAND Europe, 2009.

[252] Arianna Rossi, Rossana Ducato, Helena Haapio, Stefania Passera, and Monica Palmirani. Legal design patterns: Towards a new language for legal information design. In Erich Schweighofer et al. (Eds.), editor, *Proceedings of the 22nd International Legal Informatics Symposium IRIS 2019*, 2019.

[253] Arianna Rossi and Helena Haapio. Proactive legal design: embedding values in the design of legal artefacts. In Erich Schweighofer et al. (Eds.), editor, *Proceedings of the 22nd International Legal Informatics Symposium IRIS 2019*, 2019.

[254] Arianna Rossi and Monica Palmirani. A visualization approach for adaptive consent in the european data protection framework. In Edelmann N. Parycek P., editor, *CeDEM 2017: Proceedings of the 7th International Conference for E-Democracy and Open Government*, pages 159–170. Krems: Edition Donau-Universität Krems, 2017.

[255] Arianna Rossi and Monica Palmirani. From words to images through legal visualization. In Pagallo et al., editor, *AI Approaches to the Complexity of Legal Systems*, pages 72–85. Springer, Forthcoming.

[256] Corrado Roversi, Leonardo Pasqui, and Anna M. Borghi. *The Province of Jurisprudence Naturalized*, chapter An Experimental Study on the Grounding of Legal Concepts. Wolters Kluwer, Warsawa, 2017.

[257] Mary C Rundle. International personal data protection and digital identity management tools. presentation at igf 2006, privacy workshop i, athens, 2006, 2006.

[258] Norman Sadeh, Alessandro Acquisti, Travis D Breaux, Lorrie Faith Cranor, Aleecia M McDonald, Joel R Reidenberg, Noah A Smith, Fei Liu, N Cameron Russell, Florian Schaub, et al. The usable privacy policy project. Technical report, Technical Report, CMU-ISR-13-119, Carnegie Mellon University, 2013.

[259] Giovanni Sartor. Legislative information and the web. In *Legislative XML for the Semantic Web*, pages 11–20. Springer, 2011.

[260] Florian Schaub, Rebecca Balebako, Adam L Durity, and Lorrie Faith Cranor. A design space for effective privacy notices. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pages 1–17, 2015.

[261] Florian Schaub, Bastian Könings, and Michael Weber. Context-adaptive privacy: Leveraging context awareness to support privacy decision making. *IEEE Pervasive Computing*, 14(1):34–43, 2015.

[262] Tony Schirato and Susan Yell. *Communication and culture: An introduction*. Sage, 2000.

[263] Richard K Sherwin. *Visualizing law in the age of the digital baroque: Arabesques & entanglements*. Routledge, 2012.

[264] David Sless. What is information design. *Designing information for people*, pages 1–16, 1994.

[265] Robert H Sloan and Richard Warner. Beyond notice and choice: Privacy, norms, and consent. *J. High Tech. L.*, 14:370, 2014.

[266] Daniel J Solove. Privacy self-management and the consent dilemma. *Harward Law Review*, 126(1880), 2013.

[267] Eric R Spangenberg, Kevin E Voss, and Ayn E Crowley. Measuring the hedonic and utilitarian dimensions of attitude: A generally applicable scale. *ACR North American Advances*, 1997.

[268] Christina O Spiesel, Richard K Sherwin, and Neal Feigenson. Law in the age of images: the challenges of visual literacy. *Wagner A, Summerfield T, Benavides F (eds.) 13 Contemporary Issues of the Semiotics of Law*, pages 231–255, 2005.

[269] Helga Stevens. Tech companies aren't the enemy. Online at https://www.politico.eu/article/tech-companies-arent-the-enemy-data-scandal-facebook-mark-zucker 22 May 2018.

[270] Cass R Sunstein. Empirically informed regulation. *The University of Chicago Law Review*, 78(4):1349–1429, 2011.

[271] Cass R Sunstein. Nudging: a very short guide. *Journal of Consumer Policy*, 37(4):583–588, 2014.

[272] Cass R Sunstein. Nudging and choice architecture: Ethical considerations. *Yale Journal on Regulation*, Forthcoming.

[273] John Sweller. Cognitive load during problem solving: Effects on learning. *Cognitive science*, 12(2):257–285, 1988.

[274] John Sweller, Jeroen JG Van Merrienboer, and Fred GWC Paas. Cognitive architecture and instructional design. *Educational psychology review*, 10(3):251–296, 1998.

[275] Stefano Taddei and Bastianina Contena. Privacy, trust and control: Which relationships with online self-disclosure? *Computers in Human Behavior*, 29(3):821–826, 2013.

[276] Privacy Tech. Privacy icons. https://www.privacytech.fr/privacy-icons/.

[277] Doug Tidwell. *Xslt.* " O'Reilly Media, Inc.", 2008.

[278] Peter M Tiersma. *Legal language.* University of Chicago Press, 1999.

[279] TNS Opinion & Social. Special Eurobarometer 431 Data Protection. Technical report, Requested by European Commission, Directorate-General for Justice and Consumers and coordinated by Directorate-General for Communication, 2015.

[280] TNS Opinion & Social. Flash eurobarometer 443 eprivacy. Technical report, Commissioned by European Commission, Directorate-General for Communications Network, Content and Technology and co-ordinated by Directorate-General for Communication, December 2016.

[281] TRUSTe. Truste and disconnect introduce visual icons to help consumers understand privacy policies.

[282] UNECE. Road traffic and road signs and signals agreements and conventions. Online at http://www.unece.org/transport/international-agreements/transconventnlegalinst/list-of-agreements-for-tabs/road-traffic-and-road-signs-and-signal html.

[283] UNICEF. Un convention on the rights of the child in child-friendly language. Online at https://www.unicef.org/rightsite/files/uncrcchilldfriendlylanguage.pdf, 2006.

[284] Jeroen Van den Hoven, P Vermaas, and Ibo Van de Poel. *Handbook of ethics, values and technological design*. Springer, 2015.

[285] Mark Van Hoecke. *Law as communication*. Bloomsbury Publishing, 2002.

[286] Froukje Sleeswijk Visser, Pieter Jan Stappers, Remko Van der Lugt, and Elizabeth BN Sanders. Contextmapping: experiences from practice. *CoDesign*, 1(2):119–149, 2005.

[287] Fabio Vitali. A standard-based approach for the management of legislative documents. In *Legislative XML for the Semantic Web*, pages 35–47. Springer, 2011.

[288] W3C Policy Languages Interest Group (PLING). Platform for privacy preferences (p3p) project, 2002/2006.

[289] Anne Wagner. The rules of the road, a universal visual semiotics. *International Journal for the Semiotics of Law*, 19(3):311–324, 2006.

[290] Rob Waller. Information design: How the disciplines work together. Technical report, Simplification Centre, March 2011.

[291] Robert Waller, Jenny Waller, Helena Haapio, Gary Crag, and Sandi Morrisseau. Cooperation through clarity: Designing simplified contracts. *Journal of Strategic Contracting and Negotiation*, 2(1-2):48–68, 2016.

[292] Steffen P Walz and Sebastian Deterding. *The gameful world: Approaches, issues, applications.* Mit Press, 2015.

[293] Tim Wellhausen and Andreas Fiesser. How to write a pattern?: a rough guide for first-time pattern authors. In *Proceedings of the 16th European Conference on Pattern Languages of Programs*, page 5. ACM, 2012.

[294] Alan F Westin. Privacy and freedom. *New York atheneum*, 7, 1967.

[295] Susan Wiedenbeck. The use of icons and labels in an end user application program: an empirical study of learning and retention. *Behaviour & Information Technology*, 18(2):68–82, 1999.

[296] S. Wilson, F. Schaub, A. Dara, S. K. Cherivirala, S. Zimmeck, Mads Schaarup Andersen, P. Giovanni Leon, E. Hovy, and N. Sadeh. Demystifying privacy policies with language technologies: Progress and challenges. In *First Workshop on Text Analytics for Cybersecurity and Online Safety (TA-COS 2016)*, 2016.

[297] Shomir Wilson, Florian Schaub, Aswarth Abhilash Dara, Frederick Liu, Sushain Cherivirala, Pedro Giovanni Leon, Mads Schaarup Andersen, Sebastian Zimmeck, Kanthashree Mysore Sathyendra, N Cameron Russell, et al. The creation and analysis of a website privacy policy corpus.

In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 2016.

[298] Michael S Wogalter, Vincent C Conzola, and Tonya L Smith-Jackson. Research-based guidelines for warning design and evaluation. *Applied ergonomics*, 33(3):219–230, 2002.

[299] Michael S Wogalter, N Clayton Silver, S David Leonard, and Helen Zaikina. *Handbook of warnings*, chapter Warning symbols, pages 159–176. Lawrence Erlbaum Associates Mahwah, NJ, 2006.

[300] Jennifer Snow Wolff and Michael S Wogalter. Comprehension of pictorial symbols: Effects of context and test method. *Human factors*, 40(2):173–186, 1998.

[301] Kuang-Wen Wu, Shaio Yan Huang, David C Yen, and Irina Popova. The effect of online privacy policy on consumer privacy concern and trust. *Computers in human behavior*, 28(3):889–897, 2012.

[302] Karen Yeung. 'hypernudge': Big data as a mode of regulation by design. *Information, Communication & Society*, 20(1):118–136, 2017.

[303] Archie Zariski. *Legal Literacy: An Introduction to Legal Studies*. Athabasca University Press, 2014.

[304] Christian Zimmermann. A categorization of transparency-enhancing technologies. *arXiv preprint arXiv:1507.04914*, 2015.

[305] Frederik Zuiderveen Borgesius. Consent to behavioural targeting in european law-what are the policy implications of insights from behavioural economics? 2013.

[306] Zynga. Privacyville. Online at `https://www.zynga.com/privacy/privacyville`. Last accessed: February 7, 2018.

Appendix A

An Example of Visual Privacy Policy

Privacy Policy

Date of application: 25/05/2018

General information

This notice (the "Privacy Policy") is intended to inform you about our practices regarding the collection, use and disclosure of personal information that you may provide via this website or our mobile applications (the "Platforms").

In this notice, the following roles are described:

The controller: ....
....

The user (or data subject):...............................
....

How do we process your data?

We encrypt your data:...
....

We anonymize your data:...............................
....

We pseudonymize your data:........................
....

We make automated decisions about you: ....

We profile you:..
....

Why do we process your data?

For security purposes: ....
....

For marketing purposes: ....
....

For the provision of the service: ....
....

For purposes of service enhancement: ....
....

For research purposes:..
....

For statistical purposes: ....
....

On which legal bases?

Legitimate interest of the controller

Contract

Consent

Legal obligation

Vital interest

Public interest

Where do we keep your data?

We transfer your data outside of the EU:
....

What are your rights?

You have the right to be informed:
....

You have the right to access your data:
....

You have the right to rectify your data:
....

You have the right to restrict the processing of your data: ....
....

You have the right to erase your data:
....

You have the right to data portability:
....

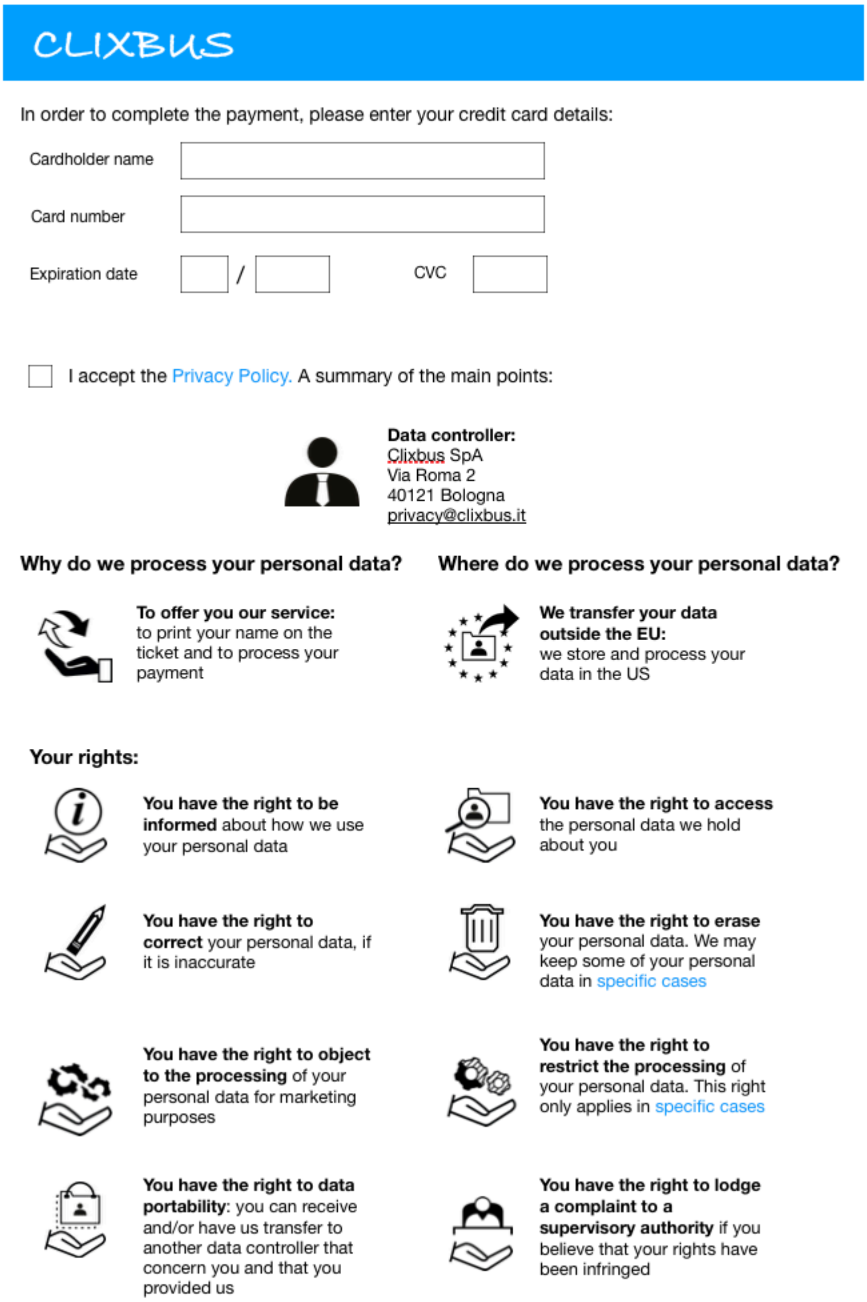
You have the right to withdraw your consent:...

You have the right to object to the processing of your data:..................................
....

You have the right to lodge a complaint to a supervisory authority: ....
....

Appendix B

The First Layer of a Multi-layered approach



CLIXBUS

In order to complete the payment, please enter your credit card details:

Cardholder name ▭

Card number ▭

Expiration date ▭ / ▭ CVC ▭

□ I accept the Privacy Policy. A summary of the main points:

Data controller:
Clixbus SpA
Via Roma 2
40121 Bologna
privacy@clixbus.it

Why do we process your personal data?

To offer you our service: to print your name on the ticket and to process your payment

Where do we process your personal data?

We transfer your data outside the EU: we store and process your data in the US

Your rights:

You have the right to be informed about how we use your personal data

You have the right to access the personal data we hold about you

You have the right to correct your personal data, if it is inaccurate

You have the right to erase your personal data. We may keep some of your personal data in specific cases

You have the right to object to the processing of your personal data for marketing purposes

You have the right to restrict the processing of your personal data. This right only applies in specific cases

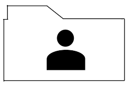
You have the right to data portability: you can receive and/or have us transfer to another data controller that concern you and that you provided us

You have the right to lodge a complaint to a supervisory authority if you believe that your rights have been infringed

Appendix C

The first DaPIS: Personal Data Types, Processing Operations, and Agents' Roles

C.1 Personal Data Types

Icon	Description	Legal ref.
	**Original personal data** Definition: it is the personal data provided by the data subject, either directly or observed from her behaviors. Rational behind the choice: Typically, folders contain data and this symbol is widespread on graphical user interfaces, whereas the user's silhouette signifies the data subject.	[13]

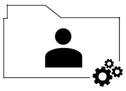
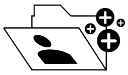
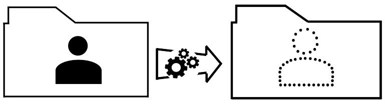
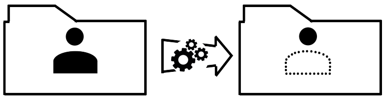
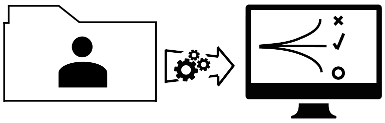
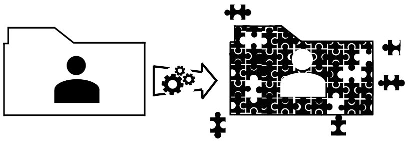
	**Processed personal data** Definition: it is the personal data after they have been processed, thus after they have been stored, organised, structured, modified,combined, etc. Rational behind the choice: Gears is a common symbol for (mechanical) transformation or processing. Used as denominator, it indicates that the data contained in the folder has been processed.	[13]
	**Derived personal data** Definition: it is the inferred and derived data generated by the controller from the analysis of the original data. Rational behind the choice: Inferred and derived data is not provided by the data subject. It originates from other data and tells something more on the data subject, hence the pluses that enter into the folder and add novel information to it.	[13]

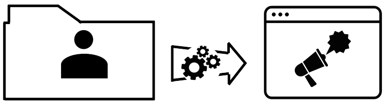
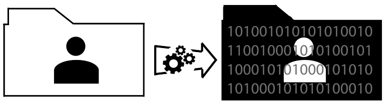
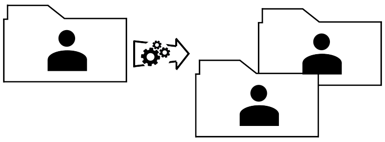
Table C.1: Icons, respective definitions, and rational behind the visual choice for the icons of the class "personal data"

C.2 Processing Operations

Icon and Description	Legal ref.
	Rec. 26 [101] ; [24]

Anonymization Definition: it is the process that strips personal data of sufficient elements such that the data subject can no longer be identified. Rational behind the choice: This icon, as the ones reported below in this table, shows a process: the personal data, on the left, are processed (represented by the arrow with gears) and become anonymous. Whereas in the icon on the left the silhouette is black to identify a specific user, it becomes blank and dotted to signify that data was striped of identifiable elements.	
 **Pseudonymization** Definition: it is the process through which personal data can no longer be attributed to a specific data subject without the use of additional information (provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person). Rational behind the choice: This icon is based on the symbol for anonymous data. The silhouette is not completely blank (='pseudo') because it is possible to re-identify the data subject, by retrieving the information that had been separated.	Art 4.5 [101]

 **Automated decision-making** Definition: it is the ability to make decisions by technological means without human involvement. Solely automated decision-making, including profiling, produces legal effects or significantly affects the data subject. Rational behind the choice: The three options stand for possible decisions that can be taken. The absence of a human, replaced by a computer, represents the fact that the decisions are taken automatically.	Art 22.1 [101]; [29]
 **Profiling** Definition: it is any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviors, location or movements. Rational behind the choice: Many pieces of a puzzle are combined together to compose the profile of a data subject.	Art. 4.4 [101]

 **Direct marketing** Definition: The communication by whatever means of any advertising or marketing material, which is carried out by the Direct Marketer itself or on its behalf and which is directed to particular individuals. Rational behind the choice: The bullhorn stands for a tool that amplifies the advertisement slogans, represented by the speech balloon. The web interface exemplifies the usual place of display of advertisements (typically online).	[104]
 101001010101010010 110010001010100101 100010101000101010 101000101010100010 **Encryption** Definition: Encryption is a mathematical function using a secret value - the key - which encodes data so that only users with access to that key can read the information. Rational behind the choice: The binary code exemplifies a digital transformation of the personal data into encrypted data that can not be read by anybody.	Art 34.3(a) [101]; [162]
	Art. 15(3), 15(4) [101]

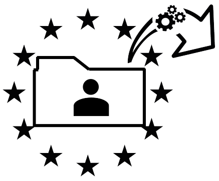
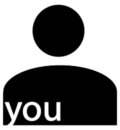
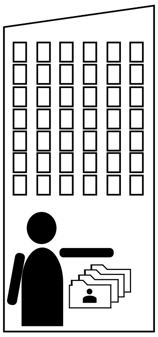
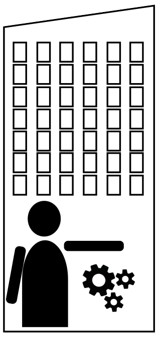
Copying Definition: It is the act of making a copy of a certain data. Rational behind the choice: Two personal data folders are exactly reproduced.	
 **Transfer of personal data to third countries** Definition: It is the transfer of personal data which are undergoing processing or are intended for processing to a third country. Rational behind the choice: The stars in circle are the emblematic symbol of the EU, whereas the arrow signifies the movement of the personal data outside of the European borders.	Art. 44 [101]

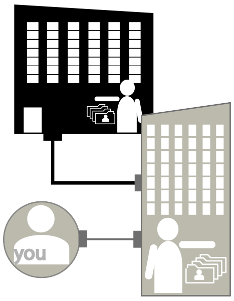
Table C.2: Icons, respective definitions, and rational behind the visual choice for the icons of the class "processing operations"

C.3 Agents' Roles

Icon	Description	Legal ref.

	**Data subject** Definition: an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person; Rational behind the choice: The data subject can be generally identified with the user of a certain service, e.g. social media. Thus the silhouette of a user widely adopted on many applications to locate one's own profile's information can easily represent the data subject. To reinforce this idea, 'you' was added to establish a direct connection with the reader.	Art. 4.1 [101]

	**Controller** Definition: it is the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data. Rational behind the choice: The controller decides the destiny of the gathered personal data, represented by the folders. For such reason, this role is symbolized by a user with one raised arm, that exercises her decision-making on the personal data. Typically, the controller is a representative of an organization, such as a company, hence the building.	Art. 4(7) [101]
	**Processor** Definition: it is the natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller Rational behind the choice: The processor's icon has a similar structure to the controller's icon, but gears are displayed under its control because it carries out the processing operations.	Art. 4(8) [101]

 	**Third party** Definition: it is the natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, is authorized to process personal data. Rational behind the choice: The third party receives the data subject's data through a controller, hence this transfer is symbolized by the cable that connects the data subject with the controller (direct transfer) and the controller with the third party (indirect transfer). The first and second parties (data subject and controller) are grayed out so that the third party can stand out.	Art. 4(10) [101]

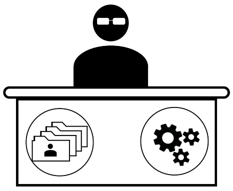
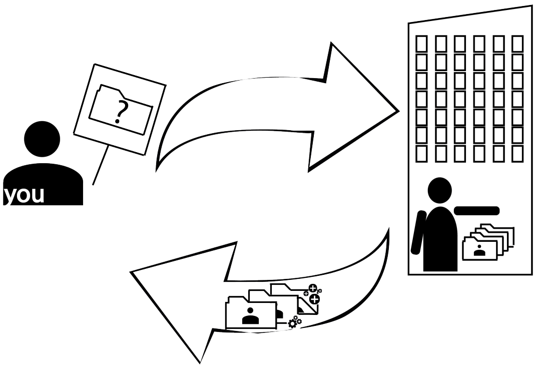
	**Supervisory authority** Definition: it is an independent public authority which is established by a Member State to be responsible for monitoring the application of this Regulation, in order to protect the fundamental rights and freedoms of natural persons in relation to processing and to facilitate the free flow of personal data within the Union. Rational behind the choice: Representing this concept as a judge would have been misleading and inherently wrong, thus the authority is sitting at a massive desk and has reading glasses to carry out analyses with the goal of ensuring that a balance between the interests of data subjects (symbolized by the data folder) and controllers/processors (symbolized by the processing gears) is respected.	Art. 4(21) [101]

Table C.3: Icons, respective definitions, and rational behind the visual choice for the icons of the class "agents' roles"

C.4 Right of Access and Right to Data Portability

Icon and Description	Legal ref.

 **Right of access** Definition: The data subject shall have the right to obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed, and, where that is the case, access to the personal data [...]. Rational behind the choice: This concept is represented as a narrative, where the data subject holds a sign on which appears a folder combined with a question mark that symbolizes the request of knowing which kind of personal data the controller has about her. The controller sends back personal data to the data subject: not only what she provided, but also the processed and inferred personal data.	Art. 15(1) [101]

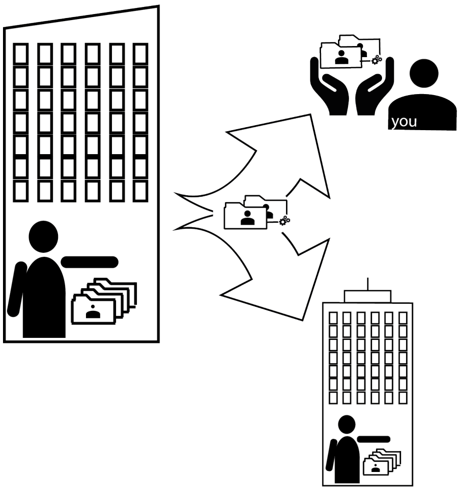
 **Right to data portability** Definition: The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided[...] [T]he data subject shall have the right to have the personal data transmitted directly from one controller to another, where technically feasible. Rational behind the choice: This concept is represented as a narrative, where the controller has two options: either she sends the (original and processed) data directly into the hands of the data subject, or to a different controller.	Art. 20 [101]

Table C.4: Icons, respective definitions, and rational behind the visual choice for the the right of access and the right to data portability

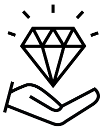
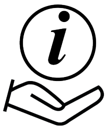
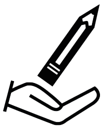
Appendix D

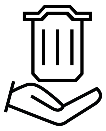
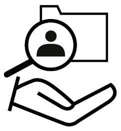
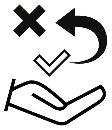
The second DaPIS: Data Subjects' Rights, Legal Bases, and Processing Purposes

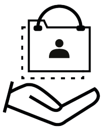
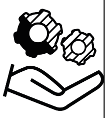
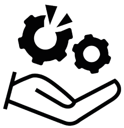
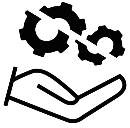
The icons are displayedin Tables D.1, D.2, and D.3,, according to the class they belong to . An English translation of the simplified definitions provided to the workshops' participants is also shown, together with the reasons behind each iconographical choice.

D.1 Data Subjects' Rights

Icon	Description	Legal ref.

	**Data subject's rights** Simplified definition: these are the rights of those (data subjects) that have provided their personal data to an organization (company, e.g. Google or institution, e.g. tax office). Rational behind the choice: the hand means "holding", with metaphorical extension "being in control" or "have the power over" to indicate the possibility granted by a right to its holder. It is an iconographical choice in common with all the other data subjects' rights, whose meaning is specified by the object above. The diamond symbolizes a value, something precious that confers some kind of power to the data subject.	Ch. 3 [101]
	**Right to be informed** Simplified definition: data subjects have the right to know who does what with their data, how, and why. Rational behind the choice: the "i" is an internationally recognized symbol for information. For the hand, see above.	Art. 12, 13,14 [101]
	**Right to rectification** Simplified definition: data subjects have the right to ask the data about them to be corrected or updated in inaccurate and complemented if incomplete. Rational behind the choice: the pencil is a widespread symbol for editing in software applications: it erases incorrect data and rewrites them correctly. For the hand, see above.	Art. 16 [101]

	**Right to erasure ('Right to be forgotten')**: Simplified definition: In some cases, data subjects have the possibility to ask for their data to be erased. Rational behind the choice: the bin is a popular symbol for erasure in software applications. For the hand, see above.	Art. 17 [101]
	**Right of access** Simplified definition: data subjects have the right to know if someone owns data about them and to obtain a copy of it. Rational behind the choice: the folder with a user's silhouette is symbol of personal data, whilst the magnifying lens on the user indicates scrutiny of a specific person's data. For the hand, see above.	Art. 15 [101]
	**Right to withdraw your consent**: Simplified definition: data subjects have the right to revoke the consent on their data processing that they had previously given Rational behind the choice: the cross ("x") and the tick ("v") derive from the representation of consent (see legal basis). The arrow goes from the tick to the cross to signal the transformation from approval/acceptance to disagreement/disapproval. For the hand, see above.	Art. 13(2c) [101]

	**Right to data portability** Simplified definition: data subjects have the right to receive a copy of their data collected by a service provider A and transfer it to a service provider B. They can also ask for direct transfer from A to B. For the hand, see above. Rational behind the choice: the data folder, representing the personal data, takes the shape of a bag with handles to carry it around.	Art. 20 [101]
	**Right to restriction of processing**: Simplified definition: data subjects have the right to ask their data to be processed exclusively for certain purposes. For the hand, see above. Rational behind the choice: the gears represent processing activities, as in other icons. The processing goes on, but only partially: half of the gears continue to work and thus are black, whereas the other half is deactivated.	Art. 18 [101]
	**Right to object to processing**: Simplified definition: data subjects have the right to ask a service to stop processing their data for a certain purpose. For the hand, see above. Rational behind the choice: the gears represent processing activities. If broken, gears stop working. Two versions were produced and it was then in the test phase determined the preferred one.	Art. 21 [101]

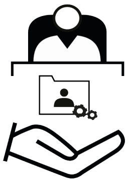
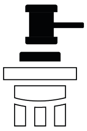
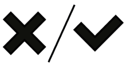
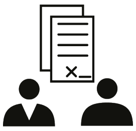
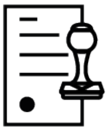
	**Right to lodge a complaint to a supervisory authority** Simplified definition: data subjects have the right to file a complaint with a supervisory authority for data protection, if they think that their data is processed unlawfully. Rational behind the choice: see supervisory authority. For the hand, see above.	Art. 13(2d) [101]

Table D.1: Icons, respective simplified definitions, and rational behind the visual choice for the icons of the class "data subjects' rights"

D.2 Legal Bases for Processing

Icon	Description	Legal ref.
	**Legal Basis** Simplified definition: It is the reason why data is processed and must be provided according to the law for the processing to be lawful. Rational behind the choice: the capital symbolizes the bases that bears the law, represented by a gavel.	Art. 6 [101]

	**Consent** Simplified definition: It is the expression of the data subject's willingness to have her data processed. Rational behind the choice: the cross ("x") represents a disagreement, whereas the tick ("v") represents an agreement. The slash conveys the idea of possibilities of an equal choice between the two, whilst the cross is expressly positioned before the tick to stress the chance to not consent, which is not usually the case, whereas the GDPR stresses the fact that consent must be freely given and informed.	Art. 6(1a), Art. 4(11) [101]
	**Contract** Simplified definition: It is an agreement that establishes a legal relationship between two parties. Rational behind the choice: The contract is usually represented as a written agreement that must be signed (hence the "x") by two parties: the data subject and the controller .	Art. 6(1b) [101]
	**Legal Obligation** Simplified definition: It is the duty to carry out what the laws says. Rational behind the choice: The law is represented as an official act, which is here signified by a stamped document with a stamp.	Art. 6(1c) [101]

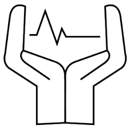
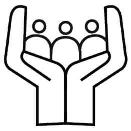
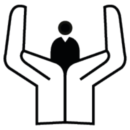
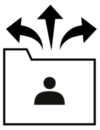
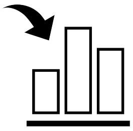
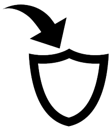
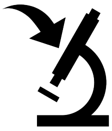
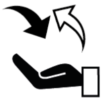
	**Vital Interest** Simplified definition: A matter of life and death. Rational behind the choice: The two joint hands stand for protection or care, with metaphorical extension for someone's interest. It is an iconographical choice in common with all the other interests, whose meaning is specified by the object between them. The electrocardiogram is an established visual convention to indicate life, as opposite to a flat tracing that means death.	Art. 6(1d) [101]
	**Public Interest** Simplified definition: It is the interest of a community, as opposed to the interest of a private. Rational behind the choice: The community is represented as three users that are blanked out, meaning that their identity is not relevant, as opposed to a specific user, which is black. For the hands' meaning, see above.	Art. 6(1e) [101]
	**Legitimate Interest** Simplified definition: It is a reason that justifies the controller's processing and that prevails on the data subject's rights. Rational behind the choice: The controller is represented as a business man. For the hands' meaning, see above.	Art. 6(1f) [101]

Table D.2: Icons, respective simplified definitions, and rational behind the visual choice for the icon for the class "legal bases for processing"

D.3 Processing Purposes

Icon	Description	Legal ref.
	**Processing Purposes** Simplified definition: the are the reasons why data is collected and processed. Without a purpose, the processing is unlawful. Rational behind the choice: this icon is a superclass of the individual purposes' classes and its iconography must be imagined together with the other purposes and the privacy policy's layout. The arrows symbolizes a direction (=a purpose): the personal data move towards a specific purpose, where the arrow lands.	Ch. 3 [101]
	**Statistical Purposes** Simplified definition: Personal data (e.g. age, gender, personal characteristics) of a certain user can be processed to carry out statistical studies on the population that the user represents. Rational behind the choice: The arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The bar graph is a typical figure to represent statistical data.	Rec. 162 [101]

	**Purposes of Information Security** Simplified definition: Personal data can be processed to ensure that the network can resist to events that can compromise its security. Rational behind the choice: The arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The shield is a common graphical symbol for security used on antivirus software and alike.	Rec. 49 [101]
	**Research Purposes** Simplified definition: Personal data can be collected and processed to carry out scientific research (e.g. medical research) Rational behind the choice: The arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The microscope is a typical and iconographical symbol for science and research.	Recc. 159, 160 [101]
	**Purposes of Provision of the Service** Simplified definition: Personal data can be processed to provide a service (e.g. Google Maps asks for user's location to provide directions). Rational behind the choice: The black arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The white, complementary arrow going in the opposite direction and departing from a black hand with a white cuff (=the controller's hand) symbolizes the service. The two arrows taken together signify the exchange of personal data for a certain service.	

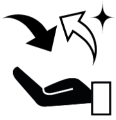
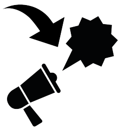
	**Purposes of Service Enhancement** Simplified definition: Personal data can be processed to enhance the functioning of a service (e.g. the navigation on a website). Rational behind the choice: Same as in provision of the service. The additional star/spark (which resembles a plus on purpose) signifies the enhancement in other digital contexts (e.g. videogames).	
	**Marketing purposes** Simplified definition: Personal data can be processed to send advertising material. Rational behind the choice: The arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The bullhorn stands for a tool that amplifies the advertisement slogans, represented by the speech balloon.	[104]

Table D.3: Icons, respective simplified definitions, and rational behind the visual choice for the icons of the class "Processing purposes"

Appendix E

The Final DaPIS

The following tables provide the icons for each data protection concept, alongside its (simplified) definition that was provided to the participants of the workshops and of the user studies and the reasons behind the iconographical choice. The last column gives indication about the legal reference from which the concept was extracted.

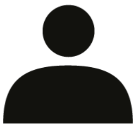
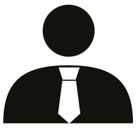
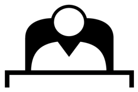
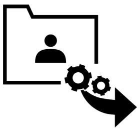
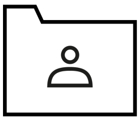
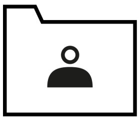
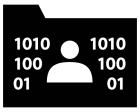
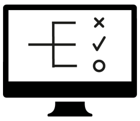
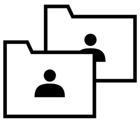
Icon	Description	Legal reference
	**Data subject** Definition: an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person; Rational behind the iconographical choice: The data subject can be generally identified with the user of a certain service, e.g. social media. Thus the silhouette of a user widely adopted on many applications to locate one's own profile's information can easily represent the data subject. To reinforce this idea, 'you' was added to establish a direct connection with the reader.	Art. 4.1 GDPR
	**Controller** Definition: it is the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data. Rational behind the iconographical choice: The controller has been one among the most debated icons and has been redesigned multiple times, based on users' feedback. Whereas in the beginning the controller was represented as a tall building (i.e. a company), and then as a man inside the building, for the sake of usability the last design iteration has given as result a business man. Indeed, it needs to combined with other elements to signify more complex notions (*see* contract, vital interest)	Art. 4.7 GDPR
	**Supervisory authority** Definition: it is an independent public authority which is established by a Member State to be responsible for monitoring the application of this Regulation, in order to protect the fundamental rights and freedoms of natural persons in relation to processing and to facilitate the free flow of personal data within the Union. Rational behind the iconographical choice: Representing this concept as a judge would have been misleading and inherently wrong, thus the authority is sitting on an armchair at a massive desk. The colour is white to distinguish it from the user and also because the user testing revealed that a black user in this context was interpreted in a negative sense (e.g. a villain)	Art. 4.21 GDPR

TABLE 1: AGENTS AND ROLES

Icon	Description	Legal reference
	**Processing operation** Definition: processing' means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction. Rational behind the iconographical choice: This icon shows the starting point of a process: the personal data, on the left, undergoes a process represented by the arrow with gears. The result is one of the icons illustrated below in this table.	Art 4.2 GDPR
	**Anonymization** Definition: it is the process that strips personal data of sufficient elements such that the data subject can no longer be identified. Rational behind the iconographical choice: The personal data, after processing, become anonymous. Whereas the icon for personal data shows a black user's silhouette to identify a specific user, here the silhouette becomes blank to signify that the data was striped of identifiable elements.	Rec. 26 GDPR; WP29, *Opinion 05/2014 on Anonymisation techniques,* 2014
	**Pseudonymization** Definition: it is the process through which personal data can no longer be attributed to a specific data subject without the use of additional information (provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person). Rational behind the iconographical choice: This icon is based on the symbol for anonymous data. The silhouette is not completely blank (='pseudo') because it is possible to re-identify the data subject, through retrieval of the information.	Art 4.5 GDPR
	**Encryption** Definition: Encryption is a mathematical function using a secret value - the key - which encodes data so that only users with access to that key can read the information. Rational behind the iconographical choice: The binary code exemplifies a digital transformation of the personal data into encrypted data that can not be read by anybody.	Art 34.3(a) GDPR; ICO Encryption (https://ico.org.uk/for-organisations/guide-to-data-protection/encryption/), 2016
	**Automated decision-making*** Definition: it is the ability to make decisions by technological means without human involvement. Solely automated decision-making, including profiling, produces legal effects or significantly affects the data subject. Rational behind the iconographical choice: The three options stand for possible decisions that can be taken. The absence of a human, replaced by a computer, represents the fact that the decisions are taken automatically.	Art 22.1 GDPR; WP29, *Guidelines on Automated Decision-making and Profiling for the Purposes of Regulation 2016/679 17/EN,* 2018
	**Copying** Definition: It is the act of making a copy of a certain data. Rational behind the iconographical choice: Two personal data folders are reproduced in an exact way.	Art. 15.3, 15.4 GDPR

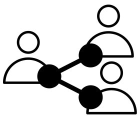
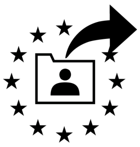
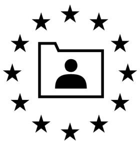
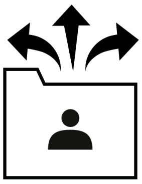
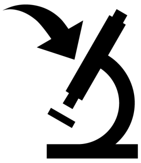
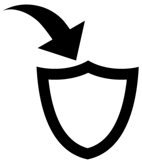
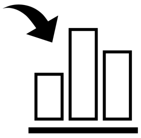
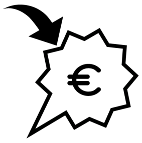
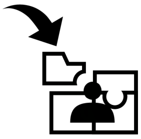
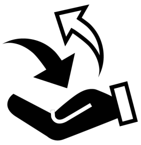
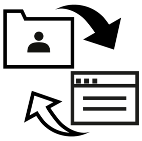
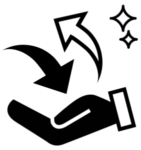
	**Sharing of personal data with third parties** Definition: It is the action of sharing personal data with a natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data. Rational behind the iconographical choice: the icon combines the ubiquitous sign for sharing, present especially on social media, with three users that are blanked out to convey the meaning that often the identity of the third parties is unknown	Art. 4.10 GDPR; in many consent forms
	**Transfer of personal data to third countries** Definition: It is the transfer of personal data which are undergoing processing or are intended for processing to a third country. Rational behind the iconographical choice: The stars in circle are the emblematic symbol of the EU, whereas the arrow signifies the movement of the personal data outside of the European borders.	Art. 44 GDPR
	**Storage of personal data in the EU** Definition: It is the opposite of the transfer of personal data to a third country. Rational behind the iconographical choice: The stars in circle are the emblematic symbol of the EU, whit a personal data folder places in the middle, i.e. inside of the European borders.	Recurrent concept in privacy policies

TABLE 2: PROCESSING OPERATIONS

Icon	Description	Legal reference
	**Processing Purposes** Simplified definition: these are the reasons why data is collected and processed. Without a purpose, the processing is unlawful. Rational behind the iconographical choice: this icon is a superclass of the individual purposes' classes and its iconography must be imagined together with the other purposes and the privacy policy's layout. The arrows symbolize a direction (i.e. a purpose): personal data move towards a specific purpose, where the arrow lands (see following icons).	Art. 6.1. GDPR
	**Research Purposes** Simplified definition: Personal data can be collected and processed to carry out scientific research (e.g. medical research) Rational behind the iconographical choice: The arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The microscope is a typical and iconographical symbol for science and research.	Recc. 159, 160 GDPR
	**Purposes of Information Security** Simplified definition: Personal data can be processed to ensure that the network can resist to events that can compromise its security. Rational behind the iconographical choice: The arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The shield is a common graphical symbol for security used on antivirus software and alike.	Rec. 49 GDPR

	**Statistical Purposes** Simplified definition: Personal data (e.g. age, gender, personal characteristics) of a certain user can be processed to carry out statistical studies on the population that the user represents. Rational behind the iconographical choice: The arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The bar graph is a typical figure to represent statistical data.	Rec. 162 GDPR
	**Marketing purposes** Simplified definition: Personal data can be processed to send advertising material. Rational behind the iconographical choice: The arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The euro symbol inside the speech balloon recalls advertisement.	Fedma, *Code of Practice for the Use of Personal Data*, 1998
	**Profiling** Definition: it is any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviors, location or movements. Rational behind the iconographical choice: The arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). Many pieces of a puzzle are combined together to compose the a personal data folder i.e. the profile of a data subject.	Art. 4.4 GDPR
	**Purposes of Provision of the Service - alternative A*** Simplified definition: Personal data can be processed to provide a service (e.g. Google Maps asks for user's location to provide directions). Rational behind the iconographical choice: The black arrow metaphorically stands for the point of arrival of the processing purpose (see superclass' icon). The white, complementary arrow going in the opposite direction and departing from a black hand with a white cuff (=the controller's hand, as opposed to the data subject's white hand in the rights icons) symbolizes the service. The two arrows considered together signify the exchange of personal data for a certain service.	Recurrent purpose in privacy policies
	**Purposes of Provision of the Service - alternative B*** Simplified definition: Personal data can be processed to provide a service (e.g. Google Maps asks for user's location to provide directions). Rational behind the iconographical choice: Since the alternative A received much criticism in the second user study because not representative of the concept, this alternative provides a more literal and semantically specified visualization: the user provides personal data in exchange of a service, exemplified by a webpage. The two arrows recall the exchange.	Recurrent purpose in privacy policies
	**Purposes of Service Enhancement - alternative A*** Simplified definition: Personal data can be processed to enhance the functioning of a service (e.g. the navigation on a website). Rational behind the iconographical choice: Same as in provision of the service (alt. A). The additional sparkling signifies enhancement in other digital contexts (e.g. videogames).	Recurrent purpose in privacy policies

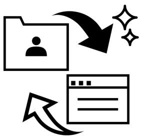
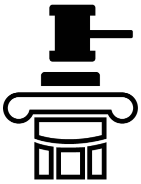
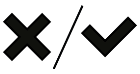
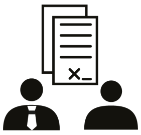
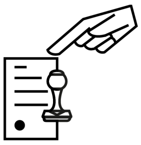
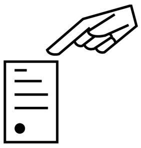
	**Purposes of Service Enhancement - alternative B*** Simplified definition: Personal data can be processed to enhance the functioning of a service (e.g. the navigation on a website). Rational behind the iconographical choice: Same as in provision of the service (alt. B). The additional sparkling star signifies enhancement in other digital contexts (e.g. videogames).	Recurrent purpose in privacy policies

TABLE 3: PURPOSES OF PROCESSING

Icon	Description	Legal reference
	**Legal Basis** Simplified definition: It is the reason why data is processed and must be provided according to the law for the processing to be lawful. Rational behind the iconographical choice: the capital symbolizes the basis that bears the law, represented by a gavel.	Art. 6 GDPR
	**Consent** Simplified definition: It is the expression of the data subject's willingness to have her data processed. Rational behind the iconographical choice: the cross ("x") represents a disagreement, whereas the tick ("v") represents an agreement. The slash conveys the idea of possibilities of an equal choice between the two, whilst the cross is expressly positioned before the tick to stress the option of refusing one own's consent in line with the GDPR, which is not usually the case.	Art. 4.11 GDPR
	**Contract** Simplified definition: It is an agreement that establishes a legal relationship between two parties. Rational behind the iconographical choice: The contract is typically represented as a written agreement that must be signed (hence the "x") by two parties: the data subject and the controller .	Art. 6.1(b) GDPR
	**Legal Obligation - alternative A*** Simplified definition: It is the duty to carry out what the laws says. Rational behind the iconographical choice: The law is represented as an official act, which is here signified by a stamped document with a stamp and a pointing hand imposed from above to recall the obligation. A previous icon design without the hand was deemed to similar to a certification and not enough legally defined.	Art. 6.1(c) GDPR
	**Legal Obligation - alternative B*** Simplified definition: It is the duty to carry out what the laws says. Rational behind the iconographical choice: Same as alternative A, but with a simplified layout for usability reasons. If the stamp is necessary to provide enough semantically defined details is still an open question.	Art. 6.1(c) GDPR

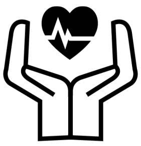
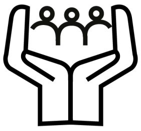
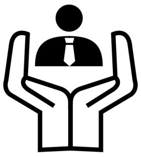
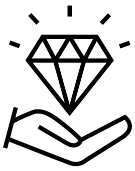
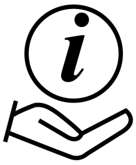
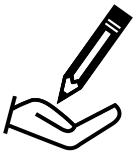
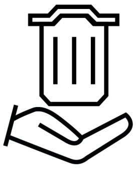
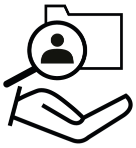
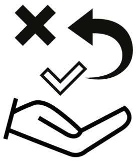
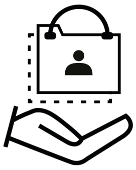
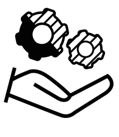
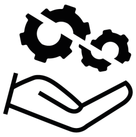
	**Vital Interest** Simplified definition: A matter of life and death. Rational behind the iconographical choice: The two joint hands stand for protection or care, with metaphorical extension for someone's interest. It is an iconographical choice in common with all the other interests, whose meaning is specified by the object between them. The electrocardiogram in the heart is an established visual convention to indicate life, as opposite to a flat tracing that means death.	Art. 6.1(d) GDPR
	**Public Interest** Simplified definition: It is the interest of a community, as opposed to the interest of a private. Rational behind the iconographical choice: The community is represented as three users that are blanked out, meaning that their identity is not relevant, as opposed to a specific user, which is black. For the hands' meaning, see above.	Art. 6.1(e) GDPR
	**Legitimate Interest** Simplified definition: It is a reason that justifies the controller's processing and that prevails on the data subject's rights. Rational behind the iconographical choice: The controller is represented as a business man. For the hands' meaning, see above.	Art. 6.1(f) GDPR

TABLE 4: LEGAL BASES FOR PROCESSING

Icon	Description	Legal Reference
	**Data subject's rights** Simplified definition: these are the rights of those that have provided their personal data (i.e. the data subjects) to an organisation (i.e. a company, e.g. Google, or an institution, e.g. tax office). Rational behind the iconographical choice: the hand means "holding", with metaphorical extension "being in control" or "have the power over" to indicate the possibility granted by a right to its holder. It is an iconographical choice in common with all the other data subjects' rights, whose meaning is specified by the element above it. The diamond symbolises a value, stressing that rights are precious and confer a certain power to the data subject.	Ch. 3 GDPR
	**Right to be informed** Simplified definition: data subjects have the right to know who does what with their data, how, and why. Rational behind the iconographical choice: the "i" is an internationally recognized symbol for information. For the hand, see above.	Art. 12, 13,14 GDPR

	**Right to rectification** Simplified definition: data subjects have the right to ask the data about them to be corrected or updated in inaccurate and complemented if incomplete. Rational behind the iconographical choice: the pencil is a widespread symbol for editing in software applications: it erases incorrect data and rewrites them correctly. For the hand, see above.	Art. 16 GDPR
	**Right to erasure ('Right to be forgotten')** Simplified definition: In some cases, data subjects have the possibility to ask for their data to be erased. Rational behind the iconographical choice: the bin is a popular symbol for erasure in software applications. For the hand, see above.	Art. 17 GDPR
	**Right of access** Simplified definition: data subjects have the right to know if someone owns data about them and to obtain a copy. Rational behind the iconographical choice: the folder with a user's silhouette is symbol of personal data, whilst the magnifying lens on the user indicates scrutiny of a specific person's data. For the hand, see above. It is the redesign of a literal representation of this concept.	Art. 15 GDPR
	**Right to withdraw your consent** Simplified definition: data subjects have the right to revoke the consent on their data processing that they had previously given Rational behind the iconographical choice: the cross ("x") and the tick ("v") derive from the representation of consent (*see* consent as legal basis). The arrow stands for the transformation from approval/acceptance to disagreement/disapproval. For the hand, see above.	Art. 13.2(c) GDPR
	**Right to data portability** Simplified definition: data subjects have the right to receive a copy of their data collected by a service provider A and transfer it to a service provider B. They can also ask for direct transfer from A to B. Rational behind the iconographical choice: the data folder, representing the personal data, takes the shape of a bag with handles to carry it around. For the hand, see above.	Art. 20 GDPR
	**Right to restriction of processing*** Simplified definition: data subjects have the right to ask their data to be processed exclusively for certain purposes. Rational behind the iconographical choice: the gears represent processing activities, as in other icons. The processing goes on, but only partially: half of the gears continue to work and thus are black, whereas the other half is deactivated. For the hand, see above.	Art. 18 GDPR
	**Right to object to processing*** Simplified definition: data subjects have the right to ask a service to stop processing their data for a certain purpose. Rational behind the iconographical choice: the gears represent processing activities. If broken, gears stop working. For the hand, see above.	Art. 21 GDPR

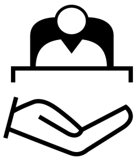
	**Right to lodge a complaint to a supervisory authority** Simplified definition: data subjects have the right to file a complaint with a supervisory authority for data protection, if they think that their data is processed unlawfully. Rational behind the iconographical choice: see icon for supervisory authority. For the hand, see above.	Art. 13.2(d) GDPR

TABLE 5: RIGHTS OF THE DATA SUBJECT

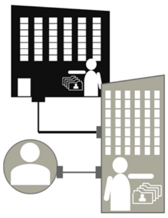
* For the icons with this symbol, research about possible alternatives that can better convey the meaning should be carried out, because in the user studies no definitive consensus was reached.

Appendix F

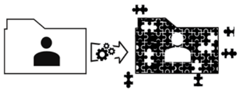
Excerpts from the First User Study

TASK 1

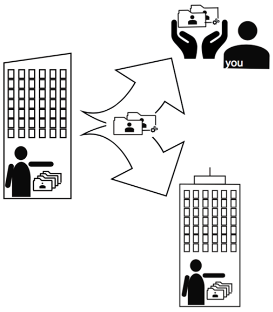
Transfer of personal data to third countries



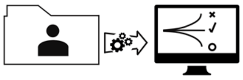
Third party



Profiling

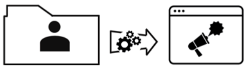


Right to data portability



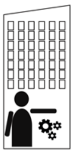
Automated decision-making

Data subject



Direct marketing

Derived personal data



Data processor

When your data is divided from you. Now you can be identified only with extra data.

This is the information about you that is transformed in some way.

This is who can receive your data, but is different from the data controller or data processor.

This is the person to whom personal data refer.

When your data is written in such a way that only authorized people can understand it.

When computers make decisions about you based on your data.

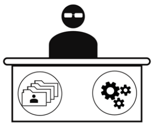
When your data is sent outside of the European Union.

This is who monitors if the law on data protection is applied and protects your rights.

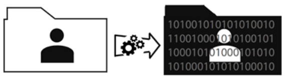
This is who collects your data and decides how your data can be processed.

You have the right to know if someone has information about you. You also have the right to receive a copy of that information.

This is the information about you that is derived from other data or by programs.

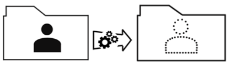


Supervisory authority



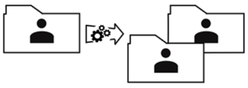
Encryption

Original personal data

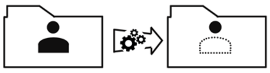


Anonymization

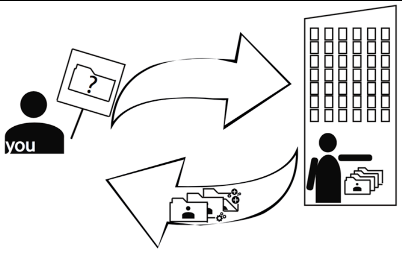
Processed personal data



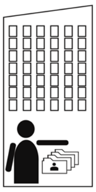
Copying



Pseudonymization



Right of access



Data controller

This is the information about you that is collected from you.

When your data is divided from you. Now you cannot be identified.

When your interests, your behaviour, or your skills are predicted based on your data.

This is who carries out operations on your data, on account of the data controller.

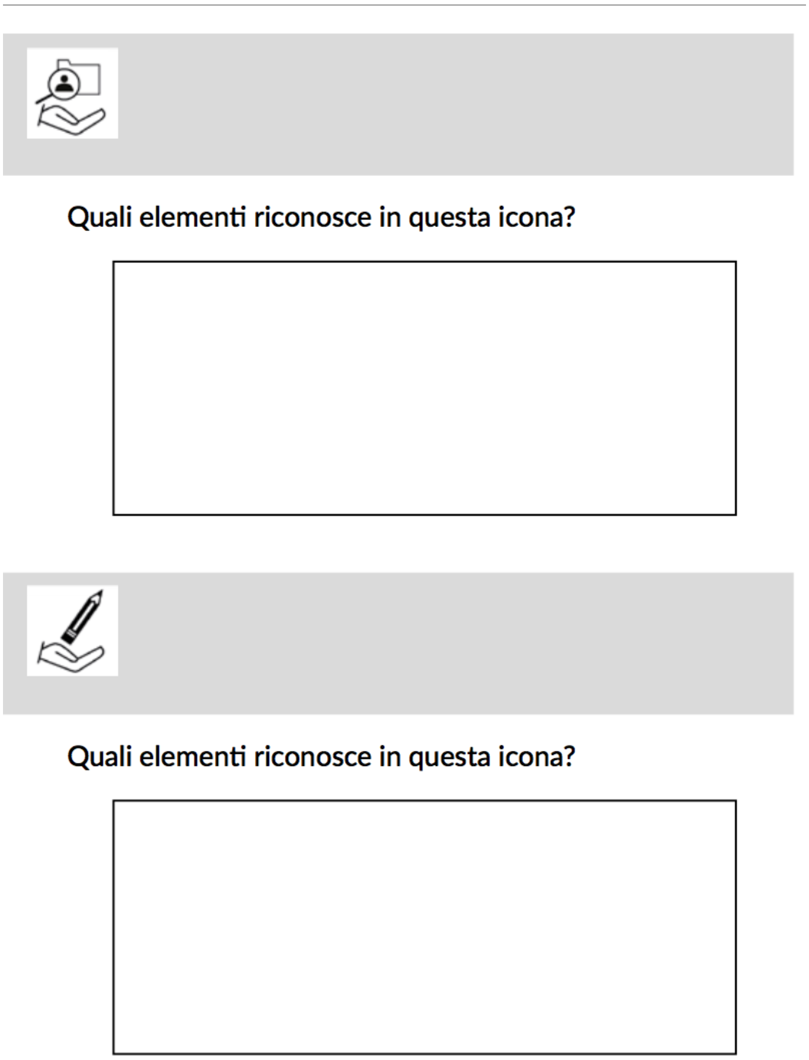
When your data is used to send you advertising.

You have the right to receive data collected about you in a format that supports re-use. You also have the right to ask the transfer of that data to another data controller.

When a copy of your data is made.

Appendix G

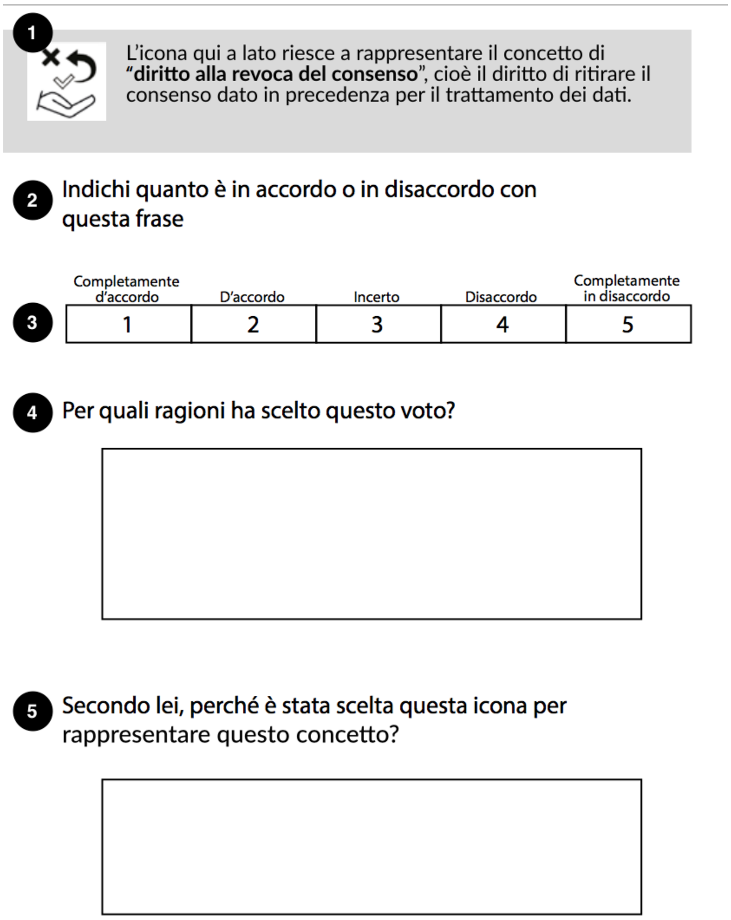
Excerpts from the Second User Study



Quali elementi riconosce in questa icona?

Quali elementi riconosce in questa icona?

Figure G.1: An example of the legibility task (task 1) in the second user study (see Sec. 6.5). The English translation would be: “Which elements do you recognize in this icon?”



1 L'icona qui a lato riesce a rappresentare il concetto di **"diritto alla revoca del consenso"**, cioè il diritto di ritirare il consenso dato in precedenza per il trattamento dei dati.

2 Indichi quanto è in accordo o in disaccordo con questa frase

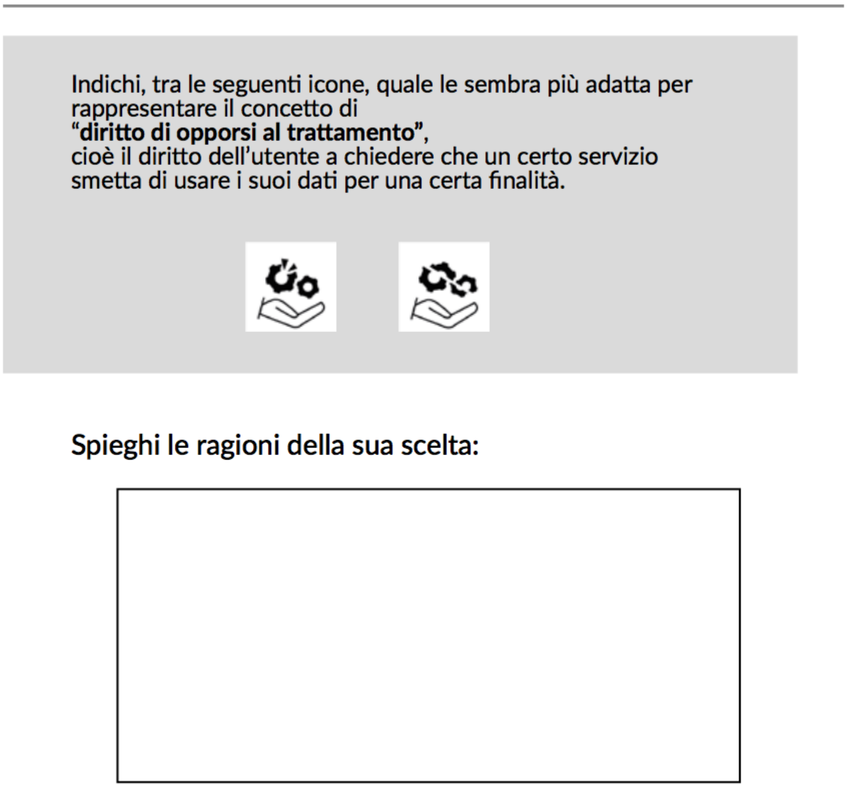
3

Completamente d'accordo	D'accordo	Incerto	Disaccordo	Completamente in disaccordo
1	2	3	4	5

4 Per quali ragioni ha scelto questo voto?

5 Secondo lei, perché è stata scelta questa icona per rappresentare questo concetto?

Figure G.2: An example from the second user study (see Sec. 6.5) of the task about ease of understanding (task 2) and the task about the alignment between designers' intentions and users' mental models (task 3). (1) displays the icon and provides the corresponding label and definition, e.g. English translation:"The icon on the side is able to represent the concept of 'right to withdraw your consent', namely the right to revoke the consent on data processing that was previously given."; (2) English translation: "Specify the extent to which you agree or disagree with this statement"; (3) 5 points Likert scale; (4) English translation: "Why have you chosen this mark?"; (5) English translation: "According to you, why was this icon chosen to represent this concept?"



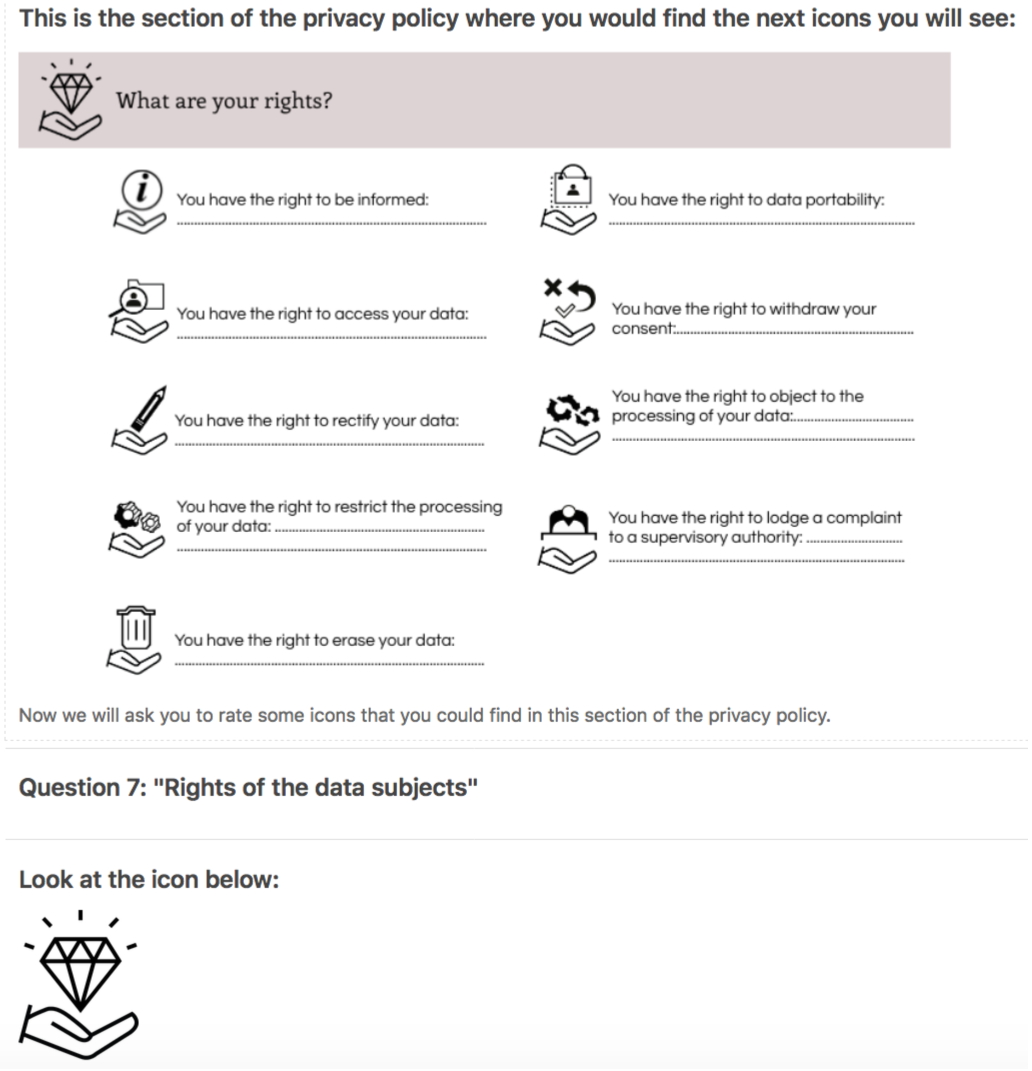
Indichi, tra le seguenti icone, quale le sembra più adatta per rappresentare il concetto di
“diritto di opporsi al trattamento”,
cioè il diritto dell’utente a chiedere che un certo servizio smetta di usare i suoi dati per una certa finalità.

Spieghi le ragioni della sua scelta:

Figure G.3: Alternative choice between two icons representing the concept of “right to object to processing” (task 4), with space to provide reasons for the choice

Appendix H

Excerpts from the Third User Study



This is the section of the privacy policy where you would find the next icons you will see:

What are your rights?

You have the right to be informed: ....

You have the right to data portability: ....

You have the right to access your data: ....

You have the right to withdraw your consent: ....

You have the right to rectify your data: ....

You have the right to object to the processing of your data: ....

You have the right to restrict the processing of your data: ....

You have the right to lodge a complaint to a supervisory authority: ....

You have the right to erase your data: ....

Now we will ask you to rate some icons that you could find in this section of the privacy policy.

Question 7: "Rights of the data subjects"

Look at the icon below:

Figure H.1: An excerpt from the third user study, that was carried out online. The questions evaluate the fitness of correspondance between icon and definition, the reasons and the possibility to align users' and designers' mental models

Appendix I

Academic Activities and Publications

This document is to certify that Miss **ARIANNA ROSSI**, enrolled, in A.Y. 2015/2016, as PhD candidate in the EM Joint Doctorate Programme in "Law, Science and Technology", has completed her doctoral studies and she will defende the PhD thesis on *2019*.

She has attended the following courses and seminars, in respect of her mobility plan.

Her mobility plan was as follows:
- First term (01/10/2015-30/03/2016) – University of Bologna;
- Second Term (01/04/2016-30/09/2016) – University of Turin;
- Third term (01/10/2016-30/03/2017) - Universitat Autonoma de Barcelona, UAB, Barcelona, Spain;
- Fourth term (01/04/2017-30/09/2017) – The University of Luxembourg;
- Third Year (01/10/2017-30/09/2018) - The University of Luxembourg.

FIRST TERM - UNIVERSITY OF BOLOGNA

BASIC COURSE (4 ECTS)
Coordinator: Prof. Monica Palmirani

- How to access to the digital resources from UNIBO.
- How to write an abstract, a poster, a paper.
- How to present in a conference.
- How to write a PhD thesis.
- How to write a scientific paper.
- Poster presentation.

ICT-LAW (3 ECTS)
Coordinator: Prof. Giovanni Sartor

- ICT of Internet
- Patent Law - IPR Law
- Privacy Issues
- e-commerce

AI&LAW (2 ECTS)
Coordinator: Prof. Antonino Rotolo

BASIC ELEMENT OF PHILOSOPHY OF LAW AND PHILOSOPHY OF SCIENCE (4 ECTS)
Coordinator: Prof. Alberto Artosi, Prof. Corrado Roversi

- The arch of knowledge
- Theories of truth
- Deduction, Induction and abduction
- The "sociological Turn" in conteporary philosophy

- Legal science

RESEARCH PROGRESS IN ARTIFICIAL INTELLIGENCE AND LAW: AN INTELLECTUAL SURVEY (1 ECTS)

Coordinator: Prof. K. Ashley, University of Pittsburgh

LATEX BASIC COURSE

Coordinator: Dr. Erica Calardo

- Basic elements
- CV
- Paper
- Thesis
- Bibliography

ONTOLOGY DESIGN PATTERN

Coordinator: Dr. Silvio Peroni

LEGAL INFORMATICS (4 ECTS)

Coordinator: Prof. Monica Palmirani

- Data, Information, knowledge web 1.0 and web 2.0 information system
- Semantic web
- XML, RDF, OWL
- Digital forencics
- e-commerce, e-government, ejustice, and open data
- Cloud forencics
- URI and naming convention
- e-commerce, e-government, ejustice, and open data
- XML and legal XML

BIOETHICS AND BIOLAW

Coordinator: Prof. Carla Faralli

- Bioethics and Biolaw
- Moral resoning in Bioethics
- Bioethics and case-law

SEMINARS

18 February 2016	Comparing legal languages and creating common/uniform terminologies
10 March 2016	Google law

SECOND TERM - UNIVERSITY OF TURIN

IPR, PRIVACY ANDA DATA PROTECTION

Coordinator: Prof. Alessandro Mantelero

- A new paradigm for data protection in the Europena Union
- Cloud computing a legal perspective

- A multidisciplinary perspective on Open data
- Patents
- The new regulation and the right to be forgotten

LEGAL INFORMATICS
Coordinator: Prof. Massimo Durante

- Online trust and the challenge of Multi-Agent System
- Comparative law Methodology, issues and perspective
- Ethics of Security and the Law
- Parternalism and rights in security context

PRINCIPLES OF COMPUTER SCIENCE
Coordinator: Prof. Guido Boella

THIRD TERM – UAB UNIVERSITAT AUTONOMA DE BARCELONA

PRIVACY ENHANCING TECHNOLOGIES (4 ECTS)
Coordinator: Prof. Antoni Roig
BIOETHICS AND BIOLAW (4 ECTS)
Coordinator: Prof. Itziar De Lecuona
SEMANTIC WEB, RELATIONAL LAW AND LEGAL ONTOLOGIES (4 ECTS)
Coordinator: Pompeu Casanovas (IDT)
AGREEMENT TECHNOLOGIES, ODR AND CROWSOURCING (4 ECTS)
Coordinator: Pablo Noriega (IIIA)

FOURTH TERM – THE UNIVERSITY OF LUXEMBOURG

The fourth term was entirely dedicated to draft and present papers, attend to conferences and to review and complete the draft of the thesis.

THIRD YEAR - THE UNIVERSITY OF LUXEMBOURG

The third year was entirely dedicated to draft and present papers, attend to conferences and to review and complete the final thesis.

Bologna, 30/09/2018

List of publications:

- Palmirani, M., Bartolini, C., Martoni, M., Robaldo, L., & Rossi, A., (forthcoming). Legal Ontology for Modelling GDPR Concepts and Norms. *Proceedings of JURIX 2018*.
- Rossi, A., & Palmirani, M., (forthcoming). From Words to Images Through Legal Visualizations. *AI Approaches to the Complexity of Legal Systems*: Springer, Berlin, Heidelberg.
- Palmirani, M., Martoni, M., Rossi, A., Bartolini, C., & Robaldo, L., (2018). PrOnto: Privacy Ontology for Legal Compliance. *Proceedings of the 18th European Conference on Digital Government ECDG 2018*. Academic Conferences and Publishing International Limited, UK, pp. 142-151.
- Palmirani, M., Martoni, M., Rossi, A., Bartolini, C., & Robaldo, L., (2018). PrOnto: Privacy Ontology for Legal Reasoning. *International Conference on Electronic Government and the Information Systems Perspective*. Springer, Cham, pp. 139-152.
- Helena, H,, Hagan, M., Palmirani, M. & Rossi, A., (2018). Legal Design Patterns for Privacy. In: Schweighofer, Eric and et al. (Eds.), *Data Protection / LegalTech Proceedings of the 21st International Legal Informatics Symposium IRIS 2018, pp. 445-450*.
- Palmirani, M., Rossi, A., Martoni, M., & Hagan, M., (2018). A Methodological Framework for the Design of a Machine-Readable Privacy Icon Set. In: Schweighofer, Eric and et al. (Eds.), *Data Protection / LegalTech Proceedings of the 21st International Legal Informatics Symposium IRIS 2018, pp. 451 - 454*.
- Rossi, A. & Palmirani, M., (2017). A Visualization Approach for Adaptive Consent in the European Data Protection Framework. In: Parycek P., Edelmann N. (Eds.), *Proceedings of the 7th International Conference for E-Democracy and Open Government (CeDEM), Krems:* Edition Donau-Universität Krems, pp. 159-170.
- Rossi, A., (2016). Representing Privacy: a Pictorial Approach. Poster and presentation at *JURIX 2016 Doctoral Consortium* (Runner-up as Best Doctoral Consortium Paper Award).
- Di Gennaro, P., Rossi, A. & Tamburini, F., (2014). The FICLIT + CS@ UniBo System at the EVALITA 2014 Sentiment Polarity Classification Task. *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 & on the Fourth International Workshop EVALITA 2014: 9-11 December 2014, Pisa*: Pisa University Press.

Other academic activities:

Summer Schools	Akoma Ntoso Summer School 2017 (http://aknschool.cirsfid.unibo.it) International Summer School LEX 2016: Managing Legal Resources in the Semantic Web (http://summerschoollex.cirsfid.unibo.it) Summer School Open Data per il Territorio: Cultura, Turismo, Ambiente 2015 (http://culta.cirsfid.unibo.it)
Certifications	By the University of California, San Diego, on Coursera: User Experience: Research & Prototyping (License CUYPBD28YJR3), Input and Interaction (License 7BD4NDS4PSNH), Human-Centered Design: an Introduction (License 3ZBKNGH4X4AK), Design Principles: an Introduction (License GT26FYXQZA24), Information Design (License XEHW32WJFS54),

Designing, Running, and Analyzing Experiments (License W6KM6Z6SKWLR),
Social Computing (License S5XXPXVD88DV).
By Michigan State University, on Coursera: Design and Make Infographics (License JTTS886HV28P).
IELTS (Score: 8.5) Certification Date Jul 2015 – Jul 2017 (License 15IT005670ROSA010A).

Organization of workshops:

- Co-organiser and member of the scientific committee for the workshop "Legal Design as Academic Discipline: Foundations, Methodology, Applications", that will be held at Groeningen, Netherlands, during JURIX 2018 (link)
- Co-organiser of the "Legal Design Workshop for the GDPR" with the CIRSFID (University of Bologna), the Legal Design Lab of Stanford University, and MIREL in March 2018 (link)**.** The workshop aimed at creating a discussion platform around transparent privacy communication among academics, policy-makers, companies technologists, designers, lawyers, etc.
- Co-organiser of the "Law & Design for Privacy" legal design workshops (link) with the CIRSFID (University of Bologna) in March 2018 for the creation of a standardised icon set to represent key concepts of the General Data Protection Regulation. Partners: Bologna Academy of Arts and Società Italiana Informatica Giuridica.

Presentations at International Conferences:

Invited speaker

- "Data Protection by Legal Design" at the CodeX FutureLaw Conference 2018, Stanford Law School, 5 April 2018
- "When Legal Design Meets the Semantic Web: Rethinking how we Interact with Data Protection" at the Legal Design Summit, Helsinki, 1 November 2017

Paper or abstract presentations

- "AI & Legal Design for the Protection of Data Subjects. Introducing the DAPRECO Knowledge Base" at Convergences du Droit et du Numerique, Bordeaux, 16 October 2018
- "DaPIS: an Ontology-based Data Protection Icon Set" at the Law via the Internet conference, Florence, 11 October 2018
- "Visualizing Legal Information: an Ontology-based Data Protection Icon Set" at the MIREL workshop 2018, within LuxLogAI (International conference on AI), Luxembourg, 17 September 2018
- "Legal Design Patterns for Privacy" & "A Methodological Framework to Design a Machine-Readable Privacy Icon Set" at the International Legal Informatics Symposium IRIS 2018, Salzburg, February 2018
- "Automatically Transposing GDPR's Requirements for Informed Consent into Visual Interfaces" at MyData 2017, Tallinn & Helsinki, August 2017
- "A Visualization Approach for Adaptive Consent in the European Data Protection Framework" at the Conference for eDemocracy and Open Government CeDEM 2017, Krems, Austria, May 2017

- “Representing Privacy: a Pictorial Approach” at the International Conference JURIX Doctoral Consortium (http://jurix2016.unice.fr/), Nice, December 2016
- “From Words to Images through Legal Visualizations” at the Workshop on Legal Knowledge and the Semantic Web (http://ekaw-lksw2016.cirsfid.unibo.it/), Bologna, November 2016

Projects:

- Visiting researcher at the Stanford Legal Design Lab in the context of the MIREL Marie Skłodowska-Curie RISE project, which aims at creating a worldwide network of academic and industrial key partners in legal informatics (http://www.mirelproject.eu/).
- Collaborator of the ALMAIDEA “Teoria sperimentale del diritto: embodied cognition e percezione del giuridico” project that introduces empirical research into the legal theory, in particular about the relationship between embodiment of legal concepts and the design of digital interfaces.
- Collaborator in the development of the Privacy Ontology (PrOnto) for the FNR/CORE DAPRECO (DAta PRotection REgulation COmpliance) project.