UNIVERSITÀ DEGLI STUDI DELL'AQUILA

DIPARTIMENTO DI INGEGNERIA E SCIENZE DELL'INFORMAZIONE E MATEMATICA

Dottorato di Ricerca in Information and Communication Technology

Curriculum: Emerging computing models:
algorithms, software architectures and intelligent systems

XXXV ciclo

Titolo della tesi

EMPOWERING USERS IN THE DIGITAL WORLD THROUGH ETHICS AND PRIVACY PROFILING

SSD
INF/01

Dottorando
Patrizio Migliarini

Coordinatore del corso
Prof. Vittorio Cortellessa

Tutor
Prof. Paola Inverardi

Co-tutor
Prof. Simone Gozzano

a.a. 2021/2022

Abstract

The thriving scope of machine ethics related problems and emerging themes about autonomous systems such as inequality, fair human-machine interaction, biasing, deception, opacity and explainability, security, unintended consequences, rights, are increasingly the focus of attention in the scientific community. The ever-increasing consideration of the interaction between human beings and computer systems made explicit by individual users as well as by political and government institutions (e.g., GDPR), is bringing to the fore the need to design ethical-enabled systems that take into account the uniqueness and peculiarities of their users and their right to be protected against the loss of privacy, autonomy, rights, individuality, morality.

This dissertation is about ethics application in the digital world and in autonomous intelligent systems, with a particular focus on human ethics analysis and synthesis to produce an ethical model to be used in automated systems. It investigates privacy as ethical dimension, surveys the literature on privacy profiling (by the meaning of gathering, analyzing, and classifying data to construct profiles of people), analyzes problems related to machine ethics and intelligent systems, with reference to critical issues on profiling, and illustrates a practical example on web security.

Specifically, it includes the following sections:

- Introduction: the vision and the need for a user empowerment in the digital world, the EXOSOUL project, the premise of privacy as ethical dimension, the dissertation and research questions overview, the contributions.

- State of the art in privacy profiling: literature review about ethical

and privacy profiling.

- Empowering users through ethical profiling: design, development, administration and results of the ethical profiling questionnaire.
- Mobile apps domain privacy profiling: an empirical study on an existing dataset collected from the fitness domain.
- Application scenario: profile driven cookies management
- Conclusions.

Some of the work presented in this dissertation expands and deepens the research published or in the process of being published together with the EXOSOUL research team, namely Prof. Paola Inverardi, Prof. Marco Autili, Prof. Davide Di Ruscio, Dr. Costanza Alfieri, Dr. Thanh Phuong Nguyen, Dr. Massimiliano Palmiero, Dr. Gian Luca Scoccia.

Acknowledgments

I want to express my gratitude to everyone who has supported, guided, or assisted me in any way. I am especially grateful to my advisor, Professor Paola Inverardi, who patiently gave me extensive instruction on conducting research, planning studies, and thinking critically. I am indebted to her for the invaluable assistance with this investigation and the exploratory work I've been doing at the University of L'Aquila. I desire to express my appreciation to the external reviewers of my thesis, Professor Radu Calinescu and Professor Ilaria Torre for the meticulous work and the precious insights, and to all the members of my doctoral committee who worked diligently alongside our research team for this happy experience that has taught me a lot, which I am really grateful for, despite the difficulties resulting from the COVID-19 pandemic and the long periods of isolation that have unfortunately ensued. I am grateful to Nicoletta, Patrizia, and Lucia, as well as all of my other loved ones, who have provided me with unwavering love and emotional support throughout my time pursuing this academic goal. I would also like to extend my gratitude to professors, researchers, and colleagues at EXOSOUL research team Costanza Alfieri, Marco Autili, Mashal Afzal Memon, Elena Belmonte, Francesca Caroccia, Simonetta D'Amico, Davide Di Ruscio, Martina De Sanctis, Donatella Donati, Simone Gozzano, Lorenzo Greco, Thanh Phuong Nguyen, Massimiliano Palmiero, Patrizio Pelliccione, Geraldina Roberti, Gian Luca Scoccia, Marco Segala, Lisandra Suarez Fernandez, Guglielmo Tamburrini, as well as all of the other individuals who were generous with their time, knowledge, ideas, and support to make it possible for me to conduct research in this fascinating field, in a friendly and collaborative environment.

Table of Contents

Chapter 1

Introduction

Users of digital devices in today's environment are always linked to the internet [70] and are being asked more often to share their own preferences in the digital world. Simple options for the device's settings, such as notification alarms, as well as pertinent ethical choices pertaining to the user's conduct, including privacy options, are included in the realm of user preferences (e.g., concerning the unauthorized disclosure and mining of personal data, as well as the access to restricted resources). All of these preferences identify the user; they are the foundation upon which her digital identity is constructed, and they will become more significant as the development of autonomous technologies and the ethical repercussions associated with them continue to advance. Much with well used applications and operating systems, the configurations that allow these preferences are frequently difficult to find and even more difficult to comprehend. Inequality, fair human-machine interaction, biasing, deception, opacity and explainability, security, unintended consequences, and rights are some of the emerging themes about autonomous systems that are garnering an increasing amount of attention in the scientific community. Additionally, the growing scope of problems related to machine ethics and emerging themes about autonomous systems are attracting more and more attention. The ever-increasing consideration of the interaction between human beings and computer systems made explicit by individual users as well as by political and government institutions (such as GDPR), is bringing to the forefront the necessity to design ethically-enabled systems that take into account the uniqueness and peculiarities of their users and their right to be protected against the loss of privacy, autonomy, rights, individuality, and morality. This is bringing to the forefront the need to design ethically-enabled systems that take into consideration

In this dissertation, we explore ethics, its formalization, and its application in autonomous intelligent systems, with a particular focus on examining and synthesizing human ethics for the purpose of creating ethical models that can be incorporated into software products and automated systems. In addition to this, it makes suggestions about software engineering and application architecture, evaluates the current state of the art in terms of machine ethics and autonomous intelligent system ethics, and carries out issue analysis in these areas.

Participants in digital platforms are continually linked to one another and to systems and are being requested more often to share their own preferences within the realm of digital technology. This may take the form of using a mobile device or being on board a (self-driving) vehicle, for example. These systems are becoming more autonomous in the sense that they may make choices independently of the users or on their behalf [94]. Recurrently their autonomy goes beyond the confines of the system and invades the prerogatives of the user. As a result of this, ethical issues, including those pertaining to privacy (for example, the unauthorized disclosure and mining of personal data, access to restricted resources), are emerging as matters of the utmost concern because they have an impact on the moral rights of each individual human being and affect the social, economic, and political spheres [2, 30, 62]. In addition to the philosophical considerations, there are two distinct ways to handle these issues, namely regulatory and technological. Europe is currently in the lead when it comes to the regulation of autonomous vehicles [6, 8], and a common EU approach to liability rules and insurance for connected and autonomous vehicles is currently being discussed [68]. Europe implemented just few years ago (May 25th, 2018) the GDPR legislation for data protection [149]. Europe is proposing efforts in addition to the scientific community and large enterprises to identify challenges and define criteria for the development of algorithms and systems that contain autonomous capabilities [5, 4, 3, 1, 67] The reality is that the digital world is slowly but surely becoming acknowledged as a potentially dangerous environment for humans. On the legislative side, new regulations will be introduced, and on the technological side, requirements for transparency and accountability will be included in software development. These activities are moving in the direction of making the digital world a more hospitable place. Regulation is essential, as is having a comprehensive understanding of the technologies involved. In spite of this, we are well aware that in actual practice, obtaining complete compliance with the regulations and standards for openness may be very challenging and may even be impossible. As stated in [22], we are confronted with a paradox: human people are acknowledged as important players and the sensitive targets; but, in the digital world, they are passive consumers, and the power and the duty to maintain their rights remain in the hands of the manufacturers of software and systems. In Floridi's "mangrove societies" metaphor [72], human people have unsecured relationships with the digital world. This leaves them vulnerable to cyberattacks. The significant obstacle, which has not been tackled up to this point, is to fully empower them. According to Inverardi [94], "there is the need to rethink the role of the various actors in the digital world by empowering the users of the digital technology both when they operate as citizens and as individuals".

1.1 EXOSOUL

In response to the research needs just mentioned, in 2019 Autili *et al.*. proposed EXOSOUL [22], an overarching research framework that aims to equip humans with an automatically generated exoskeleton, also known as a software shield that protects them and their personal data via the mediation of all interactions with the digital world that would result in unacceptable or morally wrong behaviors according to their ethical and privacy preferences. The goal of that framework is to equip humans with an exoskeleton that is automatically generated. The exoskeleton has the potential to take on a wide variety of forms, ranging from personalized soft-libraries that an individual may install on the devices that are being used to a complex software interface that an individual can “wear”, which could be ultimately be installed on a body chip. A more balanced distribution of power in the current digital environment may be achieved by providing users with the ability to customise their own exoskeletons. This will effectively place people at the center of the action. The development of exoskeletons also paves the way for unprecedented business opportunities in the field of societal-friendly applications. This is analogous to what occurred in the case of open source software, which promoted the principles of free software in opposition to the monopoly that was held by producers of proprietary software [7]. In addition, handing over some of the (digital) power to the user again helps to resolve liability concerns that arise in autonomous systems. This is achieved by transferring responsibility to users in accordance with the ethics they have selected. In point of fact, users will be protected by software exoskeletons, which will mediate their interactions with the digital world. On the other hand, users will be accountable for the repercussions of their (ethical) choices as a result of being protected by software exoskeletons, regardless of their level of understanding of the implications. The European Group on Ethics in Science and New Technologies (EGE) issued a statement on “artificial intelligence, robotics, and autonomous systems” in which it calls for an overall rethinking of the values around which the digital society is to be structured [2]. This statement serves as the primary motivation for EXOSOUL. EGE proposes “a set of basic principles and democratic prerequisites”, the most important of which is the concept of “human dignity” in the context of the digital society. Human dignity can be defined as the acknowledgment that an individual is deserving of respect in her interaction with “autonomous” technologies. A person has to have the ability to exert control over the information that is collected about them as well as the choices that are made by autonomous systems on their behalf. This has two repercussions: *i*) it transforms users from passive to active actors in the governance of their interactions with autonomous systems; and *ii*) it redefines the value and scope of system autonomy as being related to the amount and kind of respect for the users. Both of these repercussions lift users from the category of passive to active actors in the governance of their interactions with autonomous systems. In particular, *i* necessitates a software architectural perspective of the digital world that decouples autonomous systems from users as independent agents whose interaction is regulated and managed by protocols that are dependable (with respect to the user's ethical choices). This ethical-aware user-autonomous systems decoupling perspective is also advocated in Fukuyama *et al.* [79], where they propose the implementation of a *middleware* solution. As stated in [79], *middleware* is software that sits atop an existing platform and modifies the display of underlying data. It has the ability to provide consumers with more control over their data by mediating the interaction between users and platforms, enabling platforms to accommodate specific consumer preferences while rejecting the unilateral acts of dominant actors. *Middleware* providers would be obligated to disclose their products and technological capabilities so that users may make informed decisions. The platform would provide access to its servers, while the *middleware* would function as a filter to identify the content and priority of queries. This connection might take on a variety of forms, ranging from the platform curating and ranking all content to the *middleware* acting as the primary filter, with the platform serving as a complement. To stimulate the creation of a wide variety of *middleware* solutions, however, a technological foundation is necessary. This framework would offer the necessary infrastructure and incentives for *middleware* vendors to flourish, fostering competition and expanding consumers' options for data protection. In conclusion, *middleware* gives a viable answer to the issue of concentrated platform power, balancing technology's advantages with the need to protect individual privacy.

In this light, EXOSOUL, with its overarching architecture that provides for accurate, ethical user profiling and the capacity to make use of ethical reasoning and the required ethical actuators, can fall into the perspective of the *middleware* set.

Considerations of digital ethics provide assistance in addressing *ii*. Hard and soft ethics are two distinct components that are identified within the realm of digital ethics, as described in Floridi [71]. The definition and enforcement of hard ethics are delegated to law, which, although necessary, is not adequate. It does not cover the significant area of personal preferences that are specified by the term "*soft ethics*", which is the area in which the players of the digital world engage without worries at the moment. The manufacturer of digital systems must adhere to the strict ethical guidelines established by the law. However, who is responsible for taking into account the personal preferences and ideals of each individual? We contend that the user's connection with the digital environment ought to be shaped by a soft ethical framework. It is possible to turn a person into an autonomous and active user in and of the digital society by providing them with the tools to do so, such as by providing them with a software technology that supports their soft ethics. The idea of soft ethics necessitates that users be able to engage with a system in a manner that is tailored to their own preferences. This is because soft ethics reconceives the autonomy of the system as it is constrained by hard ethics. Because of this, the capacity of an autonomous system to make judgments not only has to comply with the laws, but also with the moral preferences of any user (which of course comprehend concrete preferences regarding security and privacy). These ethical choices are given more force by the exoskeleton as it links the user to the system and manages the user's connection to the digital world. It makes the ethically conscious interaction protocols for each autonomous system in the digital world a reality via its implementation. It is an architectural connection that serves as a mediator for interactions from the point of view of software design.

Part of the doctoral work consisted of the analysis and schematizing of the architecture of the exoskeleton (Fig. 1.1), breaking it down into work packages designed to elaborate and develop the sections necessary for its realization.

As depicted in figure 1.1, the architecture consists of three work packages dedicated respectively to:

- WP1: the processing and assignment of the default profile, which is derived from the user's adherence to predetermined default profiles, which are constructed by applying the state of the art in profiling. A questionnaire will be utilized to confirm compliance.

- WP2: customization and personalization of the designated default profile via direct feedback interface and user behavior analysis. This profile is used to fuel the inferential engine that synthesizes the required solution.

- WP3: System of interface between the use case and the inferential engine. As depicted in the diagram (Fig. 1.1), the entire system places the user at the center of all action, both through affirmation systems and the interface for customizing the user's ethical profile.

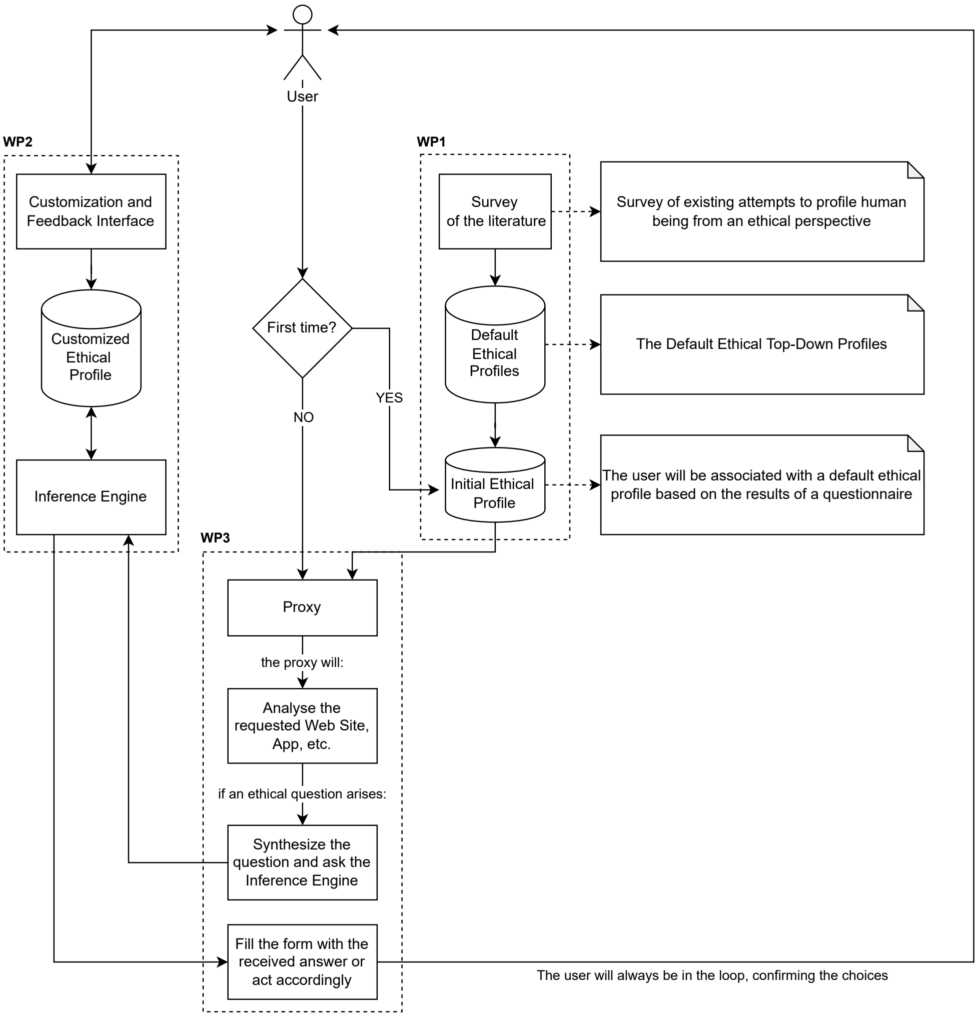

User
WP2
Customization and Feedback Interface
Customized Ethical Profile
Inference Engine
First time?
YES
NO
WP1
Survey of the literature
Default Ethical Profiles
Initial Ethical Profile
Survey of existing attempts to profile human being from an ethical perspective
The Default Ethical Top-Down Profiles
The user will be associated with a default ethical profile based on the results of a questionnaire
WP3
Proxy
the proxy will:
Analyse the requested Web Site, App, etc.
if an ethical question arises:
Synthesize the question and ask the Inference Engine
Fill the form with the received answer or act accordingly
The user will always be in the loop, confirming the choices

Figure 1.1: EXOSOUL Ethical Engine Architecture.

The subsequent dissertation focuses primarily on the theoretical investigation of potential scenarios and the implementation of some of the elements constituting the work packages just described.

In detail, relatively to:

- WP1: analysis of the literature and the construction of default ethical profiles developed in a top-down manner through a questionnaire;

- WP2: focus on a possible architecture and proposals for a recommender system;

- WP3: specific analysis of scenarios related to web interactions and the related problems of unwanted data collection and profiling.

Indeed, the general focus of this research is privacy, since we can consider privacy as an integral part of an individual's ethics. Privacy is a dimension of ethics because it is closely connected to the ethical principles of *autonomy*[88, 85], *respect for persons*[63, 25], and *non-maleficence*[73, 177], *beneficence*[97, 73], *justice*[92, 16], *veracity*[116] and more in general to *human rights*[91, 176].

Autonomy refers to the ability of individuals to make their own choices and control their own lives, and privacy is essential for protecting this ability. *Respect for persons* refers to the idea that individuals should be treated with dignity and respect, and privacy is essential for protecting this idea. *Non-maleficence* refers to the idea that individuals should not be harmed, and privacy is essential for protecting this idea. *Autonomy* is closely connected to privacy because individuals need privacy in order to make their own choices and control their own lives. Privacy allows individuals to have control over their personal information and to make decisions about how it is used. Without privacy, individuals may feel that their personal information is being collected and used without their consent or knowledge, which can lead to a loss of autonomy. *Respect for persons* is closely connected to privacy because individuals need privacy in order to be treated with dignity and respect. Privacy allows individuals to maintain control over their personal information and to make decisions about how it is used. Without privacy, individuals may feel that their personal information is being shared without their consent or knowledge, which can lead to a lack of respect for their persons. *Non-maleficence* is closely connected to privacy because individuals need privacy in order to avoid harm. Privacy allows individuals to protect their personal information and to make decisions about how it is used. Without privacy, individuals may be at risk of harm as a result of the unauthorized use of their personal information. Privacy is also closely connected to the ethical principle of *beneficence*, which refers to the idea that individuals have a moral obligation to do good and to promote the well-being of others. Privacy can promote *beneficence* by allowing individuals to control their personal information and to make decisions about how it is used, which can promote their well-being. Privacy is also closely connected to the ethical principle of *justice*, which refers to the idea that individuals should be treated fairly and that resources should be distributed in a fair and equitable manner. Privacy can promote *justice* by allowing individuals to control their personal information and to make decisions about how it is used, which can promote fair and equitable distribution of resources. Privacy is also closely connected to the ethical principle of *veracity*, which refers to the idea that individuals should be truthful and honest. Privacy can promote *veracity* by allowing individuals to control their personal information and to make decisions about how it is used, which can promote truthfulness and honesty. In addition to these ethical principles, privacy is also closely connected to the concept of *human rights*. Privacy is recognized as a human right under international law, including the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights. These documents recognize the importance of privacy in protecting human dignity and autonomy. Privacy is also recognized as a human right under international law, which further solidifies its importance in protecting human dignity and autonomy. It is important for individuals, organizations, and governments to recognize and respect privacy as a fundamental dimension of ethics in order to promote the well-being and autonomy of individuals.

Article 22 of the General Data Protection Regulation (GDPR) established the intimate connection between data protection and ethics [89]. A person has the legal right to assert that her personal information is being handled in an ethical way, according to the basic right to data protection: The ideals of human dignity and personal autonomy, both of which have a clear ethical component, are two additional moral value concepts that underpin data protection. Additionally, ethical issues are taken into account throughout the execution of data protection regulation, particularly the General Data Protection Regulation (GDPR). The application of the GDPR needs to be based, to a significant extent, on understandings of the way in which dignity, fairness, and other values are implicated, as well as on judgements about the compatibility of these situations with the realization of, for example, a liberal democratic society that is based on the rule of law and human rights. This is because threats from the "new surveillance" are becoming increasingly palpable in a rapidly changing information society in which threats from the "new surveillance" are becoming increasingly apparent. This necessitates the consideration of innovative concerns that include an ethical component. The implementation of data protection legislation is starting to include ethical analysis, which specifies the ethical substance of data protection principles and the human-rights justification of data protection law. Some of the GDPR's Recitals allude to ethical principles, such as those against discrimination and socioeconomic disadvantage, and give a framework for doing future research.

1.2 Dissertation overview

Other than the introduction and conclusion chapters, this dissertation consists of four major chapters, which are each summarized in the following paragraphs.

1.2.1 Survey of the literature about privacy profiling

Chapter 2 provides an all-encompassing examination of the work that has been done in the past on the categorization of privacy. In order to accomplish this goal, the methods that can collect data about the behaviors of individuals with regard to the privacy choices they make are being researched. In addition, various strategies for the design of privacy categories, such as Westin's segmentation, model-driven methods, and hybrid methods (e.g., data- and model-driven strategies, such as clustering + profiling and grounded analysis), are investigated, along with the ethical repercussions of using each of these strategies. The findings of the study indicate that Westin's three categories — "Unconcerned","Pragmatic" and "Fundamentalist" — should be further subdivided into a greater number of distinct categories, particularly with regard to ethical standards in the realm of digital technology. In addition, the possible upsides and downsides of each of the options are examined.

1.2.2 User's initial ethical profile: the questionnaire

The EXOSOUL exoskeleton relies on the ethical profiling of a user. This is similar in purpose to the privacy profiling that has been proposed in the literature [165, 61], but it aims at reflecting general moral preferences and predicting user's digital behaviors accordingly, like what has been proposed in [127]. To be more specific, EXOSOUL is founded on the concept of digital ethics, which can be separated into soft ethics (which reflect the ethics of the user) and hard ethics (which describe the ethical norms that a digital system must conform with). One of the goals of this research is to construct ethical profiles that can accurately predict how people will behave online. As a result, the initial stage in the EXOSOUL hybrid approach is the development of such profiles using a top-down methodology. To answer the emerging research questions, we created a questionnaire utilizing the Ethics Position Theory (EPT) [74, 77, 147], which suggests individuals' differences in moral judgments, the Five Factor Model (FFM) [51, 52] a useful framework that identifies the 5 emerging personality traits, the Hexaco theory of personality [109, 110] and dimension from the construct of Dark Triad personality [150].

1.2.3 User's ethical profile elicitation: analysis of existing data set

Through the examination of a previously collected data set, we want to get an understanding of which group of questions is most suited to distinguish users according to the privacy choices they have indicated. According to the findings, a condensed collection of semantically driven questions (relating to domain-independent privacy preferences) is a more effective method for differentiating users than a comprehensive domain-dependent one. This lends credence to the study's claim that the most important piece of information to gather is people's moral sentiments.

1.2.4 Application scenario on web data collection and cookies

In this section, we are going to examine a common circumstance in which a user is presented with the opportunity to express preferences on the effect of cookie management when they are viewing a website. We mainly focus about privacy derived from cookies management even if, in a wider perspective, we aim to take into account the communication process, including the exposure of data as well as the implications and consequences of the user's actions and decisions, as discussed in paragraph 6.2.5. In view of the GDPR directive and its characteristics, duties, procedures, stakeholders, and regulations, we are going to undertake the analysis in accordance with those provisions.

1.3 Research Questions

With this in mind, the aim of this dissertation is to answer four main research questions with corresponding sub-questions:

RQ1: *What is the state of the art in privacy categorization?*

- **RQ1.1**: *Which study contexts propose privacy categorisations that need to be identified?*
- **RQ1.2**: *How can the methodologies and approaches of privacy categorisations be understood?*
- **RQ1.3**: *How can a map be created to represent the evolution of privacy categorisations?*

RQ2: *Is it possible to elicitate user's privacy profile from existing data set?*

- **RQ2.1**: *How well does the users' self-assessment reflect their privacy category?*
- **RQ2.2**: *Which sets of questions are relevant for assessing privacy concerns?*
- **RQ2.3**: *To what extent is a recommender system able to utilize the obtained categorization in recommending relevant privacy settings to users?*

RQ3: *Is it possible to create a user's ethical profile?*

- **RQ3.1**: *Is it feasible to create a user profile that takes into account their ethical stances?*
- **RQ3.2**: *Are the ethical profiles helpful in predicting how people would behave online?*

1.4 Proposed contributions

In conclusion, this dissertation aims to provide the following contributions, including:

- Realize an extensive literature review about privacy profiling.
- Realize and administer a questionnaire that elicitates the ethical profile of a user, demonstrating it is possible to infer a default ethical profile from a top-down perspective using surveys.
- Demonstrate it is possible to select a set of general questions that can distinguish between user privacy/ethical profiles.
- Conduct a comprehensive analysis of a real-world scenario in the web domain where a software exoskeleton could empower users by reestablishing balance and preventing unwanted profiling.

Contributions will be further explored in section 6.1.

Chapter 2

State of the art in privacy categorization: a systematic review

Users in today's modern digital environment are frequently required to make decisions regarding their privacy and security that might have far-reaching repercussions. Therefore, academics are focusing more of their attention on the decisions that people make when confronted with privacy and security trade-offs, as well as the pressing and time-consuming disincentives that influence those decisions and the ways that can be used to minimize them.

While there are particular systematic studies on different elements of privacy profiling, a complete, general systematic review that offers an overview of the whole topic still needs to be completed; this is a significant gap since privacy profiling is a constantly expanding and transforming field with enormous ramifications for people, businesses, and society. A general systematic review would give a comprehensive grasp of the present state of knowledge about privacy profiling, including its definition, methodology, applications, advantages, disadvantages, and ethical and legal issues. In addition, it would emphasize areas of consensus and disagreement within the discipline, identify research gaps, and provide recommendations for future research. This systematic review seeks to address this knowledge gap by examining the present status of privacy profiling research.

The purpose of this chapter is to give a comprehensive assessment of the existing research on the classification of privacy, which has been described in terms of profiling, segmentation, clustering, and personae. This chapter examines the privacy-related behaviors of individuals and finds the positives and negatives of various privacy categorizations, as well as the ethical dimensions of these classifications. In general, the majority of the research take either a model-driven or hybrid method (for example, data- and model-driven, clustering + profiling, and grounded analysis), and it is critically based on Westin's segmentation. According to the findings of the study, the three categories that were described by Westin—namely, Unconcerned, Pragmatic, and Fundamentalist—need to be broadened in order to include more specific categories, which will finally lead to ethical classification in the digital world. There is also a discussion of the proposals' respective limitations and ramifications.

2.1 Background

Information privacy relies on the collection and use of personal data. According to Anderson (2008): «Privacy is the ability and/or right to protect your personal information and extends to the ability and/or right to prevent invasions of your personal space [...]». This definition captures both the socio-psychological perspective, which contributes to privacy-related behaviors, and the legal perspective, which poses issues related to the freedom of individuals to be protected from personal information violations and unwarranted publicity [191]. In this vein, Nissenbaum [138] refers to privacy as a product of "contextual integrity", or "socio-technical systems", in which expectations and norms regarding disclosure of information affect information flows.

Across the years, privacy has become increasingly important in people's everyday digital life, whenever they engage in online or offline activities. With the spread of the digital technology use, especially in terms of the social network services - SNSs (e.g., Facebook, Instagram, Twitter), shopping online (e.g., Amazon, Ebay), video-telephony and online chats (e.g., Zoom, Skype, Meet, Whatsapp), and remote work suites (e.g., Teams, Workspace), the understanding and the regulation of digital users' privacy protection has become a challenging area which needs to be addressed. In different terms, the use of modern systems (e.g., mobile health, financial apps, etc.) requires access to users' personal information and hosting devices, triggering not only benefits for everyone, but also privacy concerns. Indeed, information technology is being creating new social situations that challenge our assumptions about privacy and confidentiality, inevitably leading to discomfort, threat, and mistrust [24, 166].

Beyond the huge variety of different control systems for personal information protection, these concerns have motivated the adoption of regulations and laws to protect users, like the European Union's General Data Protection Regulation - GDPR [149]. However, GDPR has also posed critical issues in terms of awareness regarding the disclosure of personal information (parties that collect personal data must obtain consent from users but the legitimate interest allows to gather data without consent). Moreover, some information platforms do not offer sufficient or proper control to users, who are oriented to adopt "all-or-nothing" mechanisms in order to use the services.

This means that the concern about digital privacy does exist despite the variety of privacy control mechanisms adopted. For example, focusing on SNSs, one of the most popular control mechanisms is the "notice and choice" solution [53] that is based on the idea that users must be notified about information sharing privacy implications in order to allow them to make appropriate informed privacy decisions. Notably, such a control mechanism produced little change in users' privacy behaviors [185]. Different studies highlighted not only a lack of knowledge, given that a lot of SNS users have difficulties in managing privacy settings [112, 118], but also a lack of motivation, given that users do not fully exploit the control over their data [43]. Additionally, the "default" solution does not seem to guarantee enough protection. Default settings limits users' sharing tendency only if users show high privacy concerns [102]. Alternatively, inappropriate defaults (e.g., when the sharing information is heightened) can increase users' privacy concerns and limit the sharing behavior [192, 193].

Then, the management of the plethora of privacy options available to users can represent a problem, given that individuals have also limited cognitive resources. For example limited attention span, that does not allow them to evaluate carefully all of the conceivable alternatives and outcomes of their activities. This phenomenon of limited resources is called "bounded rationality" [170], and can prevent users' security and privacy protection. In this vein, the inherent uncertainty and ambiguity connected with the trade-offs involved in privacy and security decisions can lead the user to choose weak settings in order to unlock more functions and gain (apparently) greater value from a particular service. As discussed in Camerer *et al.*. [32], decisions involving the disclosure of information or the security of information systems are also susceptible to cognitive and behavioral biases, systematic deviations in judgements and actions from the available options of an utility-maximizing decision maker. For example, possible cognitive biases are: "anchoring", that reflects the tendency to consider information as a referent point for a specific situation (e.g., when deciding about posting on SNSs one may be affected by others" post); "framing effect", that reflects the tendency to make decisions on the basis of how options are presented (e.g., SNSs" users are more willing to disclose private information if they are offered stronger privacy controls); optimism bias and overconfidence, that refer to underestimate the possibility to get negative outcomes and to overestimate the accuracy of one's judgment, respectively (e.g., users can underestimate the efficacy of the antivirus) see [10].

Notably, also incomplete and asymmetric information may represent problems for digital privacy and security. For example, parties that manage mailing lists might sell users' information to other parties without the users' consent. This means that it is difficult for people to know the risks they are taking by using a specific system or setting, even though they know or acknowledge that their data is being gathered and exploited. The risks may occur when selecting to download an app based on its access to sensitive data, or when making a judgment on whether to trust if information has to be shared with a website configuring the browser or the cookies settings, or when opening or not a link in a document or in an email, or when answering or not a phone call from an unknown number.

Then, it is also crucial to bring out the phenomena known as the "privacy paradox" [140], which describes a situation in which there is either no correlation at all or very little correspondence between beliefs and behaviors towards privacy. This basically reflects that the instruments adopted to measure privacy (e.g., the Internet Users' Information Privacy Concerns scale [120], or the Buchanan's privacy concern and protection scale [29]) show a poor prediction power of actual digital behavior [23, 131]. Although different approaches have been proposed to solve the privacy paradox, such as the "privacy calculus" model, which focuses on how individuals share information or use privacy settings in terms of benefits and costs [55], critics emerged both at theoretical and validity levels.

Altogether these issues can lead to undesirable outcomes in terms of digital privacy and, therefore, need to be addressed. In this vein, using "soft paternalistic interventions" or "nudges", that influence decision making without altering the free will can be cosidered a valid stategy (for an extensive review see [10]. However, in order to overcome the limits related to the control mechanisms and measures adopted, and also to reduce the risks associated with the use of the digital technology, new approaches need to be developed and pursued. Amongst others users' digital privacy categorization has been used in the last decades as a prerequisite to link better privacy awareness to digital behavior, as well as to determine the extent to which the scarce knowledge of the available control mechanisms limits the users" privacy strategies and behaviors [194]. Privacy categorization involves the possibility to classify users according to specific prerequisites, such as their ability to manage privacy issues, or in terms of which type of and how many personal information they decide or do not decide to disclose. Privacy categorization has been defined and used for different purposes. For example, it has been used to both understand and support users in making choices (e.g., [19, **?**]. At the same time, it can also help in designing functionalities that improve the usability of digital technologies [59], considering both domain-specific settings (e.g., SNS, fitness, etc...) and a more general ethical setting, even though this latter has been less explored. Ultimately, as the impact of information technologies increases, and by consequence the awareness and concern about privacy enhance, the categorization in terms of actions, attitudes and values can give new insights to better understand and envisage users" digital privacy behavior. Privacy categorization can allow users to automatize decision making and choices in the digital world, still preserving their willingness. It can reflect the users" attitudes, personality, ethics, or can be based on users" cognitive processes and digital behavior.

With this in mind, the aim of this systematic review on privacy categorization is to answer three main research questions:

RQ1.1: *Which study contexts propose privacy categorisations that need to be identified?*

RQ1.2: *How can the methodologies and approaches of privacy categorisations be understood?*

RQ1.3: *How can a map be created to represent the evolution of privacy categorisations?*

2.2 Systematic Review Methodology

On May 5th, 2022, we used **5** research strings over three of the major scientific search engines (Scopus, PubMed, Web of Science) and saved all

	KEYWORDS	SCOPUS	PUBMED	WEB OF SCIENCE	TOTAL	UNIQUES
	results	**4091**	**363**	**6771**	**11225**	**6193**
1	**privacy and persona**	113	6	63	181	
2	**privacy and profiling**	648	55	2229	2930	
3	**privacy and profile**	1134	168	2229	3529	
4	**privacy and clustering**	1823	74	1988	3867	
5	**privacy and segmentation**	373	60	262	695	

Table 2.1: Keywords results per search engines with details and totals

the results. We researched the keywords in titles and abstracts obtaining **6193** unique results reported in Table 1 2.1.

The reason why 5 different search strings were used instead of a single string (e.g., constructed using the Boolean logical operators and wildcards) was because this method allowed to investigate the different distribution of the keyword *privacy* in association with different types of categorization, namely:

- *segmentation*
- *clustering*
- *profile*
- *profiling*
- *persona*

As indicated above, this work was carried out together with Prof. Paola Inverardi and Dr. Massimiliano Palmiero; we independently selected and evaluated the articles in the following stages: first by the title, then by the abstract, and lastly by the whole text, looking for those articles that were considered relevant to the study.

The eligibility criteria were the following: only articles written in English; papers focused on privacy; paper focused on proposing classifications and labels in terms of categorizations as defined above; papers focused on critics and/or evolution of the Westin's Segmentation.

Westin's Segmentation is a pioneering and influential framework for comprehending individuals' privacy concerns and preferences. This model, created by Prof. Alan Westin, established the groundwork for subsequent research in the field of privacy by categorizing individuals according to their attitudes toward privacy and personal data sharing. By capturing the nuanced differences in how individuals perceive and value privacy, Westin's Segmentation has had a significant impact on policy development, technology design, and academic research. Therefore, despite criticism from certain works, it remains essential for analyzing and understanding the complex relationship between privacy, technology, and society.

This study's methodology involved a multistage process of filtering and selecting the most pertinent academic papers for the systematic review from a large pool. Beginning with an initial set of 6,193 titles, we examined each title for relevance to the research topic. Based on this analysis, we reduced the number of papers to 43.

We then focused on the abstracts of these 43 papers, analyzing their content and evaluating their relevance to the objectives of the study. At this juncture, thirteen papers were determined to merit a full-text analysis.

After conducting an in-depth analysis of the full texts of these 13 papers, we chose six that best met the objectives of the study and provided valuable insights into the topic. These six publications were ultimately included in the systematic review, providing a sound basis for synthesizing the existing literature and deriving conclusions.

- Watson *et al.* 2015 - Mapping User Preference to Privacy Default Settings

- Dupree *et al.* 2016 - Privacy Personas: Clustering Users via Attitudes and Behaviors toward Security Practices

- Liu *et al.* 2016 - Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions

- Wisniewski *et al.* 2017 - Making privacy personal: Profiling social network users to inform privacy education and nudging

- Dupree *et al.* 2018 - A case study of using grounded analysis as a requirement engineering method: Identifying personas that specify privacy and security tool users

- Toresson *et al.* 2020 - PISA: A Privacy Impact self-assessment App Using Personas to Relate App Behavior to Risks to smartphone Users

We integrated this search result with other 18 relevant papers selected by cross-references (e.g., references analysis).

In total we selected, described and analyzed **24 papers** as reported in Table 2.2.

Table 2.2:

STUDY	DOMAIN	INSTRUMENTS	APPROACH	CLASSIFICATION ELEMENTS	SUBJECTS
Westin (1990)	Economy	Analysis + Questionnaire (4 Questions)	Model-driven	(3) General Privacy Concern Index: High, Moderate, Low	2254
Westin (1991)	Economy	Questionnaire (4 Questions)	Model-driven	(3) Consumer Privacy Concern Index: High (Fundamentalists), Moderate (Pragmatic), Low (Unconcerned)	1255
Westin (1993)	Health	Questionnaire (4 Questions)	Model-driven	(3) Medical Privacy Concern Index: High, Medium, Low	1000
Westin (1993)	Health Computing	Questionnaire (3 Questions)	Model-driven	(3) Computer Fear Index: High Computer Fear, Medum Computer Fear, Low Computer Fear	1000
Westin (1994)	Economy	Questionnaire (4 Questions)	Model-driven	(4) Distrust Index: High Distrust, Medium Distrust, Low Distrust, No Distrust	1005
Westin (1996)	Economy	Questionnaire (4 Questions)	Model-driven	(3) Privacy Concern Index: Privacy Fundamentalists, Privacy Pragmatists, Privacy Unconcerned	1005
Westin (2001)	Economy	Questionnaire (3 Questions)	Model-driven	(3) Privacy Segmentation Index: Privacy Fundamentalists, Privacy Pragmatists, Privacy Unconcerned	1529

Continued on next page

Table 2.2: (Continued)

Kumaraguru, Cranor (2005)	Economy	Literature Survey	Model-driven	(3) Privacy Segmentation, Index: Privacy Fundamentalists, Privacy Pragmatists, Privacy Unconcerned	1529
Sheehan (2002)	Internet Usage	Questionnaire (15+8 Questions)	Model-driven	(3) Privacy Segmentation Index + Typology of Internet Users: Unconcerned, Circumspect, Wary, Alarmed	889
Berendt, Gunther, Spiekermann (2005)	E-Commerce	Questionnaire (56 Questions) + Simulation	Model-driven	(4) Personal Consumer Information Cost Index: Privacy Fundamentalists, Profiling Averse, Marginally Concerned, Identity Concerned	171
Consolvo, Smith, et al. (2005)	Location Sharing	3 Phases: 1) Questionnaire and Exercises, 2) Experience Sampling Method, 3) Interview	Model-driven	(3) Privacy Segmentation Index: Privacy Fundamentalists, Privacy Pragmatists, Privacy Unconcerned + Who was requesting, Why they wanted, How user feels	16
Urban, Hoofnagle (2014)	Economy (General)	Surveys and Analysis	Data-driven and Model-driven	(2) Privacy Vulnerable, Privacy Resilient (or not useful at all?)	2203
Hoofnagle, Urban (2014)	Economy (General)	Surveys and Analysis	Data-driven and Model-driven	(3) Privacy Segmentation Index Critics: Privacy Fundamentalists, Privacy Pragmatists (to be revised), Privacy Unconcerned	2203
Woodruff, Pihur, et al. (2014)	Health Privacy (General)	Questionnaire	Model-driven	(4) Privacy Segmentation Index Critics: Fundamentalists, Pragmatists, Unconcerned + ”Fundamentalists Pragmatists”	884
Liu, Lin, Sadeh (2014)	Mobile Apps	LBE Privacy Guard Dataset Analysis	Data-driven	from 3 to 6 profiles (dataset-dependant)	4.8M
Lin, Liu , et al. (2014)	Mobile Apps	Google Play API Data Analysis	Model-driven	4 profiles: Conservatives, Unconcerned, Fence-Sitters , Advanced Users	725

Continued on next page

Table 2.2: (Continued)

Watson, Lipford, Besmer (2015)	Social Network Service: Facebook	Survey for the usage and general privacy attitudes: Westin's questions; Buchanan index; Facebook Intenisity Index	Model-driven	(3) Westin's profiles (pragmatist, fundamentalist, unconcerned) + Low, Average, High	184
Liu, Andersen, et al. (2016)	Mobile Apps	Enhanced Android Permission Manager Dataset Analysis	Data-driven	7 profiles (dataset-dependant)	84
Dupree, DeVries, et al. (2016)	General	Survey and open-ended interviews	Hybrid	5 Personas: Marginally concerned, Fundamentalists, Amateurs, Technicians, Lazy experts	32
Wisniewski, Knijnenburg, Lipford (2017)	Social Network Service: Facebook	Survey for privacy behaviours (management) and feature awareness (proficiency)	Model-driven	6 management profiles for privacy behaviour: Privacy maximizers, Self-censors, Time savers/consumers, Privacy balancers, Selective sharers, Privacy minimalists - Six proficiency profiles for feature awareness: Novices, Near-novices, Mostly novices, Some expertise, Near expertise, Experts	308
Dupree, Lank, Berry (2018)	Computer Based System: General	Open-ended interviews	Hybrid	5 personas: Marginally aware, Fundamentalists, Struggling amateurs, Technicians, Lazy experts	32
Schairer, Cheung, et al. (2019)	Health Privacy (General)	Focus Group, Interview,Questionnaire	Model-driven	(6) Philosophies of Privacy: Fatalism, Moral right, Nothing to hide, Something to hide, Personal responsibility, Tradeoff	108
Toresson, Shaker, et al. (2020)	KAUDroid dataset	PISA (privacy impact self-assessment) app	Hybrid	5 personas (descriptive)	n/a
Di Ruscio, Inverardi, et al. (2022)	General	Cross-domain dataset analysis	Model-driven	(3-4) Inattentive, Involved/Attentive, Solicitous	295

In Fig 2.1, we report an overview of the keywords found in the papers, where the size represents their frequency.

ELEMENTS	METHODOLOGY	APPROACH
Segment	Segmentation	Model-driven (may include data analysis but modeling is prevalent)
Cluster	Clustering	Data-driven (may include modeling but data analysis is prevalent)
Profile	Profiling	Hybrid (data analysis and modeling are included and are balanced)
Personae/Philosophies	Personification	Hybrid (data and model are included, grounded analysis is added)

Table 2.3: Elements, methodologies and approaches that constitutes *privacy categorization*

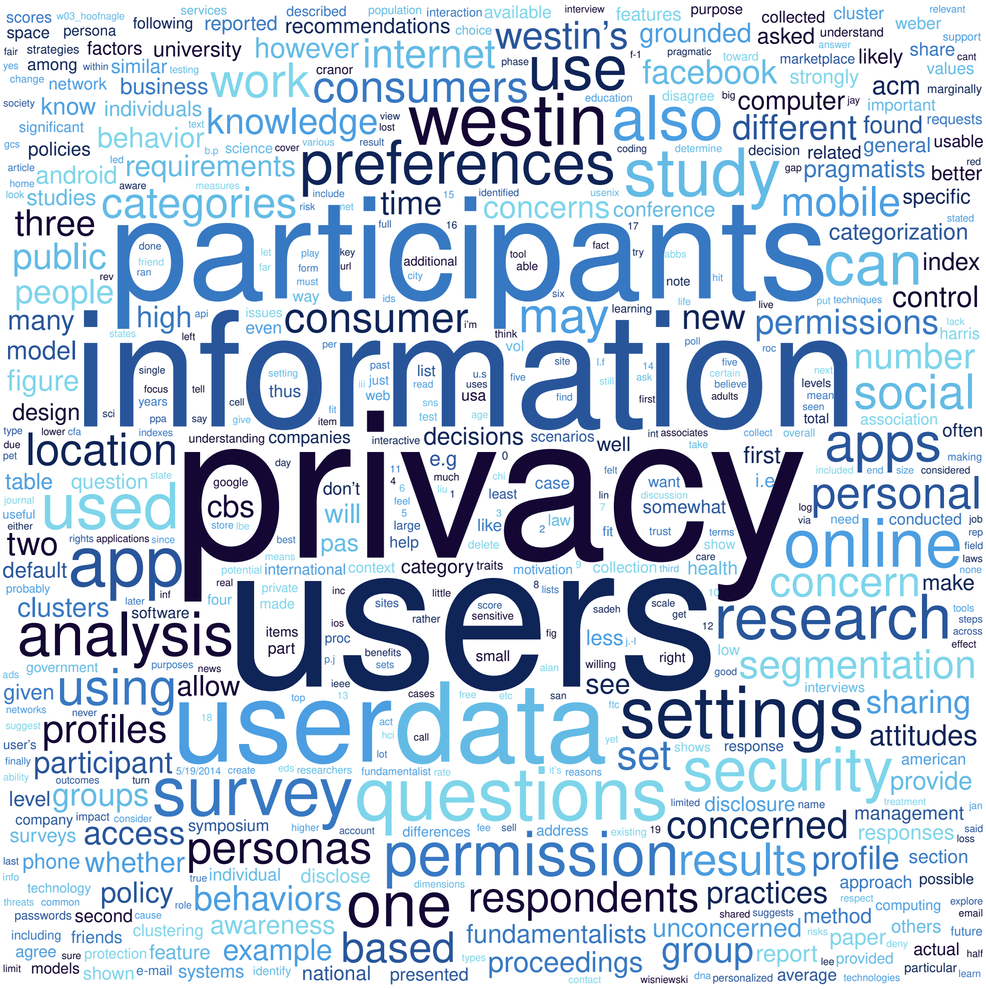


Figure 2.1: Overview of the keywords found in the papers analyzing the whole text body, shown with size relative to frequency

2.3 Basic concepts and definitions

During the literature review we observed three different approaches to privacy categorisation. We refer to those approaches as:

- **Model-Driven**
- **Data-Driven**
- **Hybrid**

The primary focus of a **Model-driven** approach is the creation and use of domain models, which are conceptual representations of the subjects relevant to a particular issue. Instead of focusing on the actual data, it places more attention on and pursues abstract representations of the knowledge and activities that regulate a specific application domain.

Taking an approach that is **Data-driven** involves basing all choices and procedures on the information provided by the data. A research that is data-driven indicates that the choices are based on the analysis and interpretation of data. A data driven method gives researchers the ability to study and organize their data with the intention of gaining a deeper grasp of the facts without the introduction of bias due to the researcher's own experience or to existing theories.

In addition to the two approaches already mentioned, we have repeatedly observed the use of a **Hybrid** approach, which contemplates the application of both a data-driven and a model-driven part of research - often in that order - which adds to a strictly data-driven analysis, further interpretative proposals guided by models emerging from the data. This hybrid approach can also incorporate **Grounded analysis** based on the grounded theory methodologies [121], in which inductive reasoning is used in the process of developing the analysis. This technique stands in contrast to the hypothesized-based, deductive approach that is typical of conventional scientific investigation.

We indicated and reported these differences and the papers descriptions in Table 2.2.

Based on the approaches just described, different methodologies are applied to obtain different categorizations:

- **Segmentation** is the process of partitioning a data set into meaningful regions or extracting relevant features from the data set [196]; this process is mainly model-driven and is used to create segments.

- **Clustering** is the process of mathematically grouping similar objects into different groups [119]; it is mainly data-driven, although it can include also modeling, and it is used to create clusters.

- **Profile/Profiling** involves mostly hybrid approaches. Specifically, it relies on correlated data created with different methodologies in order to identify and represent a subject (individual or group) [90, 159]. The correlated data aggregation occurs from different sources and individuals are usually not aware of this process. The group profiling process can be distributive (e.g., the same characteristics apply to both the group and all its members) or non-distributive (the attributes of the group do not apply to all the members and the association is statistical rather than determinate) [159].

- **Persona**, researchers usually utilize a hybrid approach (personification), by inducing and attributing further parameters to existing segments or clusters. This means that personification makes use of data and model-driven methods in conjunction with grounded analysis [60]. Indeed, a digital persona is "a model of an individual's public personality based on data and maintained by transactions, and intended for use as a proxy for the individual" [36]. Notably, it is possible to distinguish between a projected persona (e.g., a personal page on a website) and an imposed persona (e.g, identity formed by data by third parties) [36]. A digital persona involves a specific representation which is generally used only for a given context (e.g., social networks). In addition, although the individual is aware of the digital persona, she might not be aware of the all data represented, especially in the case of imposed persona [159].

Regarding the instruments, most of the studies used questionnaires to collect data about privacy behaviors or privacy permissions. The dataset analysis, based also on previous studies, and interviews defined by open questions were also adopted, whereas other instruments such as focus group, self-assessment of apps, simulation and literature review were sparsely used (Fig.2.3).

2.4 Review of existing categorizations

In this section we review the theories and methodologies of privacy preferences profiling, proceeding in chronological order and developing the analysis with respect to the evolution and criticisms to the more established theories.

As reported in Fig. 2.5, the birth of the modern privacy profiling starts with Westin's studies and evolves through critics and revisions mainly of the *Pragmatists* group he proposed. The original studies could not take into account the modern plethora of problems regarding the privacy management of the continuously online modern life that Floridi describes as *Onlife* [70], but the founding principles were revised across the years to be adapted to the needs of the digital society.

2.4.1 Westin's methodology (1970-2003): birth and evolution of the Westin's segmentation

2.4.1.a Westin's short biography

Alan Furman Westin (1929-2013) was an Emeritus Professor of Public Law & Government at the Columbia University. He was the former publisher of Privacy & American Business, and former President of the Center for Social Legal Research. As a consumer survey expert - mostly for Herris-Equifax in the marketing field - he consulted on more than 100 consumer surveys over his career, covering from general privacy to consumer privacy, medical privacy, and other privacy-related areas. His well-known privacy segmentation technique is frequently employed in a broad range of applications. Despite the fact that Westin was also a prominent historian and professor of privacy legislation, his survey research grew out of his work as a consultant to information-intensive companies [171] and he did not publish it in academic publications. As a result, it has only been subjected to a few in-depth examinations, all of which seem to have gone unaddressed by the author [82, 81, 106].

2.4.1.b Westin's segmentation

Since its inception, the Westin's segmentation has been utilized by academics from a wide range of areas to conduct privacy analyses. For example, it has been used in psychology, marketing research, computer security, and information and communications technology settings. Beyond academy, it is acknowledged that segmentation has also had a significant impact on privacy regulation in the United States [78, 9], where it serves as the foundation for the dominant "notice and choice" regime, under which consumers are expected to make informed decisions about products and services based on their personal preferences after being informed about privacy trade-offs. Essentially, the "notice and choice" model argues that customers will behave as "privacy pragmatists", and that privacy fundamentalists" preferences are strong enough to police the marketplace and sway consumers who are less active [93].

According to the original 1990/1991 work, Westin's privacy segmentation [188], people can be divided into three groups: *Privacy Fundamentalists*, *Privacy Pragmatists*, *Privacy Unconcerned.*

2.4.1.c Westin's privacy indexes

To create subdivision into groups, he created and used multiple *privacy indexes.* They evolved through the years and we report here some milestones from [188, 190, 106].

The "General Privacy Concern Index" that was established as part of Westin's research in 1990 was the first privacy index. Westin utilized a series of four questions to divide respondents into three groups, each of which represented a different degree of privacy concern, with the goal of gaining a deeper understanding of the issues of privacy. According to the findings of the study conducted by Westin in 1991 [190], the following questions were posed to respondents:

1. «Whether they are very concerned about threats to their personal privacy today.»
2. «Whether they agree strongly that business organizations seek excessively personal information from consumers.»
3. «Whether they agree strongly that the Federal government since Watergate is still invading the citizen's privacy.»
4. «Whether they agree that consumers have lost all control over circulation of their information.»

The responses to these questions were used to categorize each respondent into one of the following groups based on their level of privacy concern:

- High: 3 or 4 answers that express privacy-concern
- Moderate: 2 answers that express privacy-concern
- Low: 1 or no answers that express privacy-concern

In this context, "answers that express privacy-concern" refers to responses that demonstrate a strong interest in privacy, highlighting a preference for personal data protection and an awareness of potential privacy risks.

Although based on the questions and principles listed above, thus considering privacy as an ethical value that can be abstracted from the specific domain, Westin proposes indices that have been adapted and renamed according to the specific application cases, allowing for their timely use on the specific case study domain under consideration:

- The Equifax Report on Consumers in the Information Age (1990): *General Privacy Concern Index*
- Harris-Equifax Consumer Privacy Survey (1991): *Consumer Privacy Concern Index*

- Health Information Privacy Survey (1993): *Medical Privacy Concern Index*

- Consumer Privacy Concerns (1993): *Computer Fear Index*

- Equifax-Harris Consumer Privacy Report (1994): *Distrust Index*

- Equifax-Harris Consumer Privacy Report (1996): *Privacy Concern Index*

2.4.1.d Latest Westin segmentation categories

In 2002 [187], Westin provided the most comprehensive summation of the three categories that is available today that he named *Privacy Segmentation* Index:

- «*Privacy Fundamentalists* (about 25% of the national public): This group sees privacy as an especially high value, rejects the claims of many organizations to need or be entitled to get personal information for their business or governmental programs, thinks more individuals should simply refuse to give out information they are asked for, and favors enactment of strong federal and state laws to secure privacy rights and control organizational discretion.»

- «*Privacy Unconcerned* (about 20%): This group doesn't know what the "privacy fuss" is all about, supports the benefits of most organizational programs over warnings about privacy abuse, has little problem with supplying their personal information to government authorities or businesses, and sees no need for creating another government bureaucracy (a "Federal Big Brother") to protect someone's privacy.»

- «*Privacy Pragmatists* (about 55%): This group weighs the value to them and society of various business or government programs calling for personal information, examines the relevance and social pro-

> priety of the information sought, wants to know the potential risks to the privacy or security of their information, looks to see whether fair information practices are being widely enough observed, and then decides whether they will agree or disagree with specific information activities, with their trust in the particular industry or company involved a critical decisional factor. The Pragmatists favor voluntary standards and consumer choice overlegislation and government enforcement. But they will back legislation when they think not enough is being done - or meaningfully done - by voluntary means.»

Between 1990 and 2003, Westin's segmentation was applied based on the construction of three categories with differing privacy concerns (High, Medium, and Low). This is illustrated by a summary table in Kumaraguru's work [106]. Those groups are the basis for the development of the various indices created (e. g. Consumer Privacy Concern Index, Medical Sensitivity Index, Distrust Index, etc.). The group in the middle (Medium/Pragmatists) is always the most populated, and this uneven population has been the basis for the works that has critiqued, reworked and expanded Westin's work in the following years.

2.4.1.e Kumaraguru and Cranor (2005)

Kumaraguru and Cranor [106] presented a report to help researchers in gaining an understanding of Westin's work. They only provided the results obtained by Westin, not performing any survey or study to evaluate the values presented. They also showed that most of the indices created by Westin cannot be directly compared. They described that in his surveys, Westin developed a number of privacy indices to summarize his findings and to demonstrate trends in privacy concerns among the general population. They also summarized the Westin's surveys results as described in the end of the previous paragraph. Kumaraguru and Cranor reported how Westin has utilized a procedure to get the indexes (e.g., 1990 [190] and 1996 [189] studies) which they claim to be incorrect since the results produced by these different surveys cannot be directly compared. The indices obtained in the different studies did not use the same criteria (questions) and because the options (answers) used for obtaining the indexes were different across studies it was impossible to compare them. Westin also did not construct or offer procedures or comparison criteria for the indices he used for his surveys [188], which would have allowed a more direct comparison.

They also propose a summary of the different aspects that Westin used for deriving the privacy indices:

- General Privacy Concern Index (1990): Whether they are very concerned about threats to their personal privacy today. Whether they agree strongly that business organizations seek excessively personal information from consumers. Whether they agree strongly that the Federal government since Watergate is still invading the citizen's privacy. Whether they agree that consumers have lost all control over circulation of their information.

- Consumer Privacy Concern Index (1991): Agreement for the statements: - Consumers have lost all control over how personal information about them is circulated and used by companies.
  - My privacy rights as a consumer in credit reporting are adequately protected today by law and business practices.

- Medical Privacy Concern Index (1993): Ever used the services of a psychologist, psychiatrist, or other mental health professional. Do you believe your personal information has been disclosed? And there were other 4 questions which were all related to medical information.

- Computer Fear Index (1993): If privacy is to be preserved, the use of computers must be sharply restricted in the future. Concern level in usage of computers in medical services (patient billing, accounting).

- Distrust Index (1994): Technology has almost gotten out of control Government can generally be trusted to look after our interests. The way one votes has no effect on what the government does In general business helps us more than harm us.

- Privacy Segmentation and Core Privacy Orientation Index (1995-2003): Consumers have lost all control over how personal information is collected and used by companies. Most businesses handle the personal information they collect about consumers in a proper and confidential way. Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today.

2.4.2 Evolution, Critics, and Departures from Westin's Segmentation

Listed below are the works we found during our research that took their cues from Westin, critiquing it, extending it or using it in contexts other than the original one, from mobile applications to health. These works are reported below in chronological order of appearance. The ones that draw their line of research directly from Westin's work are followed by a specific comment through quotation marks. The others although do not explicitly derive from Westin's workare still based on categorization. We want to highlight that (with the exception of *Hoofnagle* et al. *(2014)* [93] which theoretically analyzes Westin's work) all the works found by this research pertain to the digital world.

Sheehan (2002) [169] researched privacy concerns connected to internet communication, computer use, and demographic data in consumers. Four distinct online consumer groups based on privacy concerns were created. An email poll was issued to 3724 people whose email addresses were created by a directory search engine to gauge current views about privacy concerns. The 889 completed surveys gave a nationally representative sample of internet users. 15 random statements were provided. The respondents were asked to analyze each statement from the perspective of a personal (as opposed to commercial) user of the Internet. This was followed by a 7-point bipolar scale with 1 (not at all bothered) to 7 (very concerned) being the end points. It turns out that the study's respondents are more pragmatic than those in Westin's research. Westin's typology was updated to categorize customers into four groups. The pragmatists were separated into two groups so this segmentation show two pragmatisms. First, there was a group of respondents whose overall worry was somewhat greater than the indifferent group of customers. Second, there is a group of customers that are worried but not as much as the most concerned group. Thus, four categories were created: *unconcerned, circumspect, wary, and alarmed* .

«Given these findings that indicate that each of the four groups exhibit unique characteristics, it is apparent that Westin's tripartite typology is too limited to use for categorizing online users.» [169]

Berendt et al. *(2005)* [26] investigate privacy concerns related to online shopping, focusing on actual self-disclosing behavior and the impact of privacy statements. Four privacy profiles were created by an internet purchasing experiment. 206 Participants completed a survey before buying. More than a quarter of the questions asked about respondents' readiness to divulge private data, faith in privacy declarations, importance of privacy, and responses to different privacy circumstances. Then, participants were invited to buy cameras and clothing online, which were discounted by 60% off local shop pricing. An anthropomorphic shopping bot helped with the purchase. Those participants who chose to purchase had to pay for it. Then, a PCIC (personal consumer information cost) index that took into account the validity and relevance of each response in the sales environment, as well as the difficulty of responding it was formed. A PCIC of 0 suggests the user can answer the question truthfully. A high PCIC indicates consumers are hesitant to provide this information. PCIC is substantially connected with legitimacy and relevance, and somewhat correlated with difficulty, according to regression analysis. The four groups are: *Privacy Fundamentalists, Profiling Averse, Marginally Concerned, Identity Concerned.* In particular, privacy fundamentalists and marginally concerned are more worried about giving personal information like their name, email, and postal address, whereas “profiling averse” users are more concerned about sharing personal information like their interests, hobbies, and health condition.

«A clustering of the answers to privacy-related questions revealed four different groups of users. We clearly distinguished a group of privacy fundamentalists and another group of only marginally concerned users as found elsewhere» [26]

Consolvo (2005) [44] tried to better understand the consumers’ value when deciding whether or not to share the current location with others. A three-phased research was run with 16 non-technical volunteers in the Seattle region in July 2004. Phase 1 looked at participants” social networks and how they expected to utilize location-enhanced computing. In Phase 2, the Experience Sampling Method (ESM) was used to capture participants” responses to hypothetical requests for their location from people on the buddy lists created in Phase 1. Then, in Phase 3, participants remarked on their experiences in the research and their perspectives about location-enhanced computing. The Westin/Harris Privacy Segmentation Model was used in Phase 1 to determine which basic privacy group each participant belonged to: fundamentalist, pragmatist, or unconcerned. Important factors were “who” was asking (including details like the participant’s present attitudes toward them), “why” they needed to know where the participant was, “what” knowledge would be most valuable to them, and “how I feel”. In addition, participants either revealed the most helpful (but not necessarily the most thorough) information about their location or did not disclose it at all. User location and activity (Adams’ context) was found to be a factor of lesser importance. Most notable findings are that: 1) participants want to disclose what they think would be useful to the requester or deny the request; 2) participants’ privacy classification, as determined by the Westin/Harris Privacy Segmentation Model, was not a good predictor of how users would respond to requests for their location from social relations.

«participants' privacy classification, as determined by the Westin/Harris Privacy Segmentation Model, was not a good predictor of how they would respond to requests for their location from social relations» [44]

Urban et al. *(2014)* [183] explained Westin's privacy segmentation, conducting a textual analysis, and giving empirical results, both calling into question long-held assumptions about consumer's privacy knowledge and preferences, and shedding fresh light on customers" privacy knowledge and preferences. They contradict Westin's idea that people" marketplace privacy decisions are conscious and intentional. Rather, the majority of customers have significant gaps in their awareness of privacy legislation and corporate operations. Authors critic also Westin's classification of customers as fundamentalists, pragmatists, or unconcerned. Kumaraguru and Cranor discovered that Westin often asked respondents questions that were categorized as "privacy pragmatists" or "all other respondents". Urban *et al.* claimed that this approach is problematic from a logical standpoint, since pragmatism demands believers to conduct positive inquiry, consider the costs and advantages of various choices, and reject idealism in favor of realistic methods and attainable aims. Based on this, they conducted two surveys on privacy knowledge and issues. Both polls were national, telephonic (landline and wireless) surveys. The 2009 study surveyed 1,000 Internet users, and the 2012 study surveyed 1,203. From the results, the authors proposed the *privacy vulnerable* and *privacy resilient* groups drawn from their knowledge and attitude instruments are more usable segmentation categories than the tripartite Westin's categories. In conclusion, the authors say that there is further work to be done in order to understand whether the categories are usable, and advance doubts about the segmentation approach.

«The segmentation instrument does not establish that individuals are pragmatic in theory or in practice» [183]

Hoofnagle et al. *(2014)* [93] contributed to the ongoing debate about legacy Westin's privacy segmentation model. To assess customer privacy attitudes and choices in the marketplace, Westin's privacy segmentation paradigm is widely used. It promotes a romantic and optimistic picture of customers as independent, logical, and deliberate individuals who influence company models via their own marketplace judgments. The framework Westin envisioned for privacy regulation would allow a realistic homo economicus to defend himself in the marketplace. A vast group of American customers is labeled as "pragmatists" by Westin's privacy segmentation model without any evidence that they really participate in pragmatism-defining debates.

Since 2009, the Berkeley Consumer Privacy Survey has been used by researchers, policymakers, businesses, and other stakeholders interested in understanding consumer attitudes and behaviors related to privacy and to gather data from a series of countrywide consumer polls investigating Americans' knowledge and perspective on information flows and privacy [93]. Whenever they've used Professor Westin's privacy segmentation questions, they've paired them with other assessments of knowledge and attitudes to see how well they hold up. It has been found throughout the Berkeley survey series that American consumers take measures to protect their privacy, are skeptical of businesses that require them to give up personal information in exchange for a service, wish for legal privacy rights that do not currently exist, and, most importantly, appear to operate in the marketplace with a "knowledge gap" concerning existing legal protections and actual business practices. They were able to pick up on this idea in the Berkeley Consumer Privacy Survey. One consumer survey was conducted in 2009, another in 2012, and three more in 2013. 58 Each poll used Westin's three screening questions to divide respondents into three groups—pragmatists, fundamentalists, and the unconcerned—before going on to test consumers" familiarity with and opinions on a wide range of topics that evolved with the market. Finally, they mapped these customers into Westin's privacy segmentation to evaluate its efficacy for a few of the queries.

The empirical study also shows that many customers negotiate privacy choices based on basic misconceptions about company operations, privacy regulations, and data usage limits, which may cause them to assume greater protection than is really available. When provided with true market-based information privacy options, most people want greater power than they now have. These misconceptions undermine the privacy market by leading customers to assume they don't need to bargain for privacy. As a whole, their data reveal that many people make bad choices concerning business and legal safeguards, and that Westin "pragmatists" know less than either "fundamentalists" or the "unconcerned". Finally, contrary to Westin's depiction, pragmatists join fundamentalists in rejecting information-intensive service alternatives. Thus, they believe that the most widely referenced component of Westin's work—his definition of consumer choices as pragmatic—should be seriously questioned. Westin's approach referred to users as "super-consumers." It conflated purposeful choice with the fact that most customers must accept offered business models. Users as super-customers undermines our view of the privacy market and blames consumers for the rise of privacy-invasive services. This shifts the policy discussion from market structures to consumers. The market segmentation approach and rational choice should be reviewed. It is proposed to create tools that accurately depict consumer knowledge and preferences to design successful policies. A more responsive and effective consumer protection system would encourage consumers to learn more and better align their legal protection expectations with reality. Instead, customers are kept in the dark about which technical services are socially acceptable. The loss of Professor Westin's key work on privacy and segmentation model would be a pity, rather, they propose to focus on his core thesis that people, not machines, can and should make choices regarding information flows and data privacy as machine processing becomes more complex. Westin's notion of people working as homines economici when making privacy choices has waned, but his urge for society to make these judgments has not. It is porposed to reexamine Westin's "Privacy and Freedom" teachings in light of today's privacy issues.

«Further, our empirical research supports and goes beyond more general experimental work to reveal that many consumers negotiate privacy preferences based on fundamental misunderstandings about business practices, privacy protections, and restrictions upon the use of data, and that these misunderstandings may lead them to expect more protection than actually exists» [93].

Lin et al. *(2014)* [111] identified privacy profiles that can help users in managing their mobile app privacy preferences. Four privacy profiles were created using the hierarchical clustering approach (Canberra distance and average linkage agglomerative method). At the aim, 725 participants' privacy preference ratings of 837 free mobile apps downloaded by Google Play, reflecting people's comfort with the purpose for which different apps request their permissions, were collected. Static code analysis was used to infer the purpose associated with a given app's permission. In addition, different types of 3rd-party libraries responsible for requesting access to a given permission were distinguished. Results showed that user's willingness to grant permissions depends on the purpose associated with the permissions, and that all mobile apps privacy preferences cannot be captured by default settings. Most importantly, the clustering approach showed the following privacy profiles:

- *privacy conservatives* (lack of comfort granting permissions; uncomfortable with mobile apps asking to access phone ID, contact list or SMS functionality);

- *unconcerned* (high level of comfort disclosing sensitive personal data, with the exception of granting SNS libraries access to the Get_Accounts permission (e.g., information linked to Facebook, Goolge+, Youtube; in general they are younger and with lower education);

- *fence-sitters* (in the between of the extremes, being quite comfortable disclosing sensitive personal data; similar to pragmatists);

- *advanced users* (high nuanced understanding of which usage scenarios they should be concerned about, e.g., they dislike targeted ads and mobile analytic libraries but agree to disclose coarse location; in general, they are older and with higher level of education).

Liu et al. *(2014)* [115] analyzed people's privacy preferences with respect to permissions to different mobile apps. Three to six privacy profiles were created using a k-means algorithm. Each user was modeled as a 12-dimensional vector of app-permission decisions (1 = allow; -1 = deny), with profiles relying on single permissions (5) and permission pairs (5) with the highest discriminating scores. Specifically, privacy preferences were analyzed as users granted permissions to different mobile apps. 4.8 million smartphone users of a mobile security and privacy platform (LBE Privacy Guard Dataset) were asked about 12 privacy settings of mobile apps downloaded. Profiles were defined according to precision, interpretability and understandability, and stability. According to the findings of this research, it is feasible to dramatically minimize user burden while still enabling consumers to have more control over their mobile app permissions. The researchers demonstrated, in particular, that simple tailored classifiers might be developed to anticipate a user's app permission choices.

Woodruff et al. *(2014)* [197] studied the links between general privacy attitudes (such as the Westin Privacy Segmentation Index), reactions to hypothetical scenarios, responses to outcomes, personality factors, and demographics. Authors conducted a two-phase research on Amazon's Mechanical Turk, which provided full data from 884 participants in order to study the association between Westin Privacy Segmentation Index beliefs and concrete effects. They also used Google Consumer Questionnaires to perform supplementary surveys (GCS). They designed a two-phase study. Phase 1 consisted of a survey that included several measures of general privacy attitudes. They aimed to capture a wide range of concerns about online and/or offline contexts. Phase 1 also included questions designed to measure participants' degree of direct and/or indirect experience with misuse of personal information. Finally, Phase 1 assessed personality characteristics using scales from the psychology literature, specifically, they included: TIPI (Ten Item Personality Inventory); locus of control; MFT (Moral Foundation Theory); general disclosiveness (subscales amount, depth and honesty); generalized self-efficacy; SIRI (Stimulating-Instrumental Risk Inventory); ambiguity tolerance; hyperbolic discounting; and CRT (Cognitive Reflection Test). Phase 2 consisted in administering a set of 20 scenarios to the same participants, who were asked to imagine themselves in three (out of 20) randomly chosen scenarios. Results showed a contradiction between participants' general privacy views and their actual or planned privacy-related activities. In addition, they performed another 3-questions-poll to see how the Westin Privacy Segmentation Index relates to participants" answers to five hypothetical circumstances. The result was that the attitude-behavior dichotomy and the attitude-consequence dichotomy were not found consistent with the Westin Privacy Segmentation Index individual items or derived categories. Authors concluded that it is possible that broad privacy attitudes represent underlying preferences that are not completely accounted for by contextual or practical concerns.

«We did not find evidence that either the individual questions or the derived categories of the Westin Privacy Segmentation Index are predictive of either participants' behavioral intent or their reaction to specific consequences, suggestive of both an attitude-behavior dichotomy and an attitude-consequence
dichotomy.» [197]

Watson et al. *(2015)* [186] explored the extent to which default privacy settings on social network sites (Facebook) are more customized to the preferences of users. Three different segmentation models were used in order to determine if multiple canonical policies are useful to improve default settings of Facebook users. Privacy profile preferences reflecting 29 profile items were collected from 184 Facebook users by a survey. Usage and general privacy attitudes toward the alternate setting options were also measured by: 3 Westin's questions, the Buchanan *et al.*.'s (2007) privacy concern questions, and the Facebook Intensity Index questions, incorporating usage and emotional connectedness to the site and its integration into people's daily activities. Firstly, from the privacy profile preferences an optimal default policy was generated for a training set of participants and was compared with the completely restrictive policies: the preferred audience largely chosen, and the permissive Facebook default settings. Secondly, from the usage and general privacy attitudes three privacy segmentation models were derived: Westin/Harris' model (pragmatist, fundamentalist and unconcerned), and Buchanan's and Facebook Intensity Index models, by which participants were divided in low, average and high according to the standard deviation from the means (the average group ranged from -1 to +1 standard deviation, whereas the low and high groups were below -1 and above +1 standard deviations from the means, respectively). Then, these models were used to determine if multiple canonical policies improved default settings: no improvement of policy fit within the data was found.

Liu et al. *(2016)* [114] created privacy profiles for permission settings and used these profiles in a personalized privacy assistant, aimed at supporting configuration of permission settings. Seven privacy profiles were created using a hierarchical (agglomerative) clustering approach. 84 participants were asked to report their privacy permission settings. Data were aggregated in terms of app category, permission and purpose associated with each permission, and for each cell a value reflecting the tendency of the user to allow or deny permissions requested by apps from a specific category when a corresponding purpose was defined. In light of the profiles identified by Lin *et al.*. (2014) [111], results showed that profiles 1, 2, 5, 6, and 7 aligned to fence-sitter" and "advanced users" profiles; profile 3 corresponded to "unconcerned" profile; profile 4 corresponded to "conservative" profile. Then, the profiles built were used to evaluate the effectiveness and usability of the profile-based personalized privacy assistant for mobile app permissions. 72 users (different from the previous field-study) were used (49 treatment group and 23 control group). Results showed that 78.7% of the recommendations made by the personalized privacy assistant were accepted, whereas only 5.1% of the recommendations were revised in a second moment by participants in the treatment group as compared to the control group. In addition, the treatment group converged faster on their settings and were also satisfied with recommendations and the personalized privacy assistant.

Wisniewski et al. *(2017)* [194] profiled Facebook users both in terms of feature awareness and privacy behavior, and explored the relationships between users' privacy awareness and behavior, in order to understand if users' privacy management strategies are affected mostly by conscious behaviors or by the limited knowledge of the available privacy controls. Six behavioural profiles of privacy management strategies and six awareness profiles for privacy proficiency were created using the mixture factor analysis. This approach allows to assign each participant to one of K classes, minimizing the residual difference between the observed and predicted factor scores for each participant. Using 308 Facebook users, both privacy behavior and feature awareness were measured by a self-report questionnaire focused on both the settings adopted to manage interpersonal privacy boundaries (e.g., I did not provide this information to Facebook; How often have you done the following to modify posts on your News Feed), and the proficiency related to a specific interface feature or functionality useful for a task (e.g., I vaguely recall seeing this item). Thus, firstly, confirmatory factor analyses were performed in order to determine the dimensional structure of both privacy behaviour and feature awareness, leading to eleven and six dimensions, respectively. Then, the mixture factor analysis was applied to the confirmed factors, leading to a six-class solution for privacy behaviour (management profiles):

- *privacy maximizers* (higher levels of privacy across the most of privacy features);
- *self-censors* (infrequent use of privacy features and settings, but high withhold of personal information);
- *time savers/consumers* (similar to privacy minimalists, but also passive consumption of Facebook updates, such as restriction of chat availability);

- *privacy balancers* (moderate levels of privacy management behaviours);
- *selective sharers* (advanced privacy settings, such as creation of friend lists and post content selectively to these groups);
- *privacy minimalists* (fewer privacy strategies, such as limiting Facebook profile by default);

and a six-class solution for feature awareness (proficiency profiles), which basically varied in degree, from the most basic to the highest level:

- *novices*;
- *near-novices*;
- *mostly novices*;
- *some expertise*;
- *near-expertise*;
- *experts.*

In general, there are some overlaps between privacy management profiles and privacy proficiency profiles: privacy maximizers are experts or near-experts; self-censors and time savers/consumers exhibit intermediate levels of proficiency; privacy balancers can show the higher or intermediate levels of expertise, or can be complete novices; selective sharers also show higher levels of expertise; privacy minimalists range from mostly novices to complete novices.

Dupree et al. *(2016-2018)* [60, 61] categorized the user space of the privacy and security features of the computer based system. Five personas were created using the grounded analysis. This latter relies on the application of inductive reasoning: ideas/concepts emerge from the data and are tagged with codes, which in turn - as the data collection progresses - can be grouped into higher-level concepts and categories which are the basis of a new theory. In details, 10 steps of grounded analysis were applied to 32 users: 1) question preparation using documents and materials made available by users; 2) conducting open-ended interviews on knowledge and concerns about privacy and security, specific security (e.g., password schemes) and protection (e.g., use of social networking sites) practices, assessment of vulnerability and invulnerability, and rational for practices and personal assessments; 3) transcription of any single thought, called *quotation*; 4) open coding of *quotations*; 5) clustering quotations with the same code to form concepts; 6) constructing an affinity diagram among concepts and sub-concepts; 7) axial coding of concepts to reveal the causal relationships among the concepts; 8) exposing properties for categorization to understand if quotations fall into properties; 9) clustering participants based on quotations; 10) creating personas, characterized by a name, photo, personal, demographic and domain specific information, a quote, and a profile of traits related to the interaction with the computer based system. Therefore, firstly users were analyzed using Westin's categorization, (pragmatist, fundamentalist and unconcerned). By this procedure two dimensions emerged and were used to describe participants, namely knowledge and motivation, especially with respect to the pragmatist category. Then, in the second categorization, users' observed similarities in *quotations* were combined with the grading of participants along knowledge and motivation, leading to five personas:

- *marginally aware* (low knowledge and motivation);
- *fundamentalist* (high knowledge and motivation);
- *struggling amateur* (medium knowledge and motivation);
- *technician* (medium knowledge and high motivation);
- *lazy expert* (high knowledge and low motivation).

Regarding Facebook's current privacy and security controls, this categorization was compared with Westin's segmentation in order to determine the extent to which groups of users showed difficulty placing themselves in only one category. In general, the five persona categorization was found to cover better the user space than the Westin's segmentation (see Dupree *et al.*., 2016).

«Overall, we posit that the domain of user segmentation would continue to benefit from additional inquiry to verify whether this five-way segmentation is an artifact of the contemporary privacy context (e.g., social networking, mobile computing, pervasive connectivity, terrorism and security, etc.) or if it is, indeed, a more accurate segmentation than Westin's prior categories.» [60]

Schairer et al. *(2019)* [165] highlighted a model of privacy dispositions by a qualitative research on privacy considerations in the field of emerging health technologies. Six philosophies of privacy were created. According to the authors, a model of privacy disposition emerged from qualitative study on privacy issues in emerging health technology. The study included 108 people in 44 interviews and 9 focus groups to examine how people value (or do not value) control over their health information. Transcripts of interviews and focus groups were coded and analyzed in ATLAS.ti for respondents' privacy concerns. Three significant qualitative results lead to a privacy disposition model. First, individuals defined their privacy-related actions as both contextual and habitual. Then there are the factors that influence the decision to share personal data that do not fall into the risk/benefit analysis. Third, privacy ideologies, or attitudes towards privacy, that are either motivational or dissuasive. A basic but powerful conceptual model of privacy disposition, or the factors that contribute to a person's overall level of privacy. It is possible to conduct more quantitative research since the components of privacy disposition have been operationalized and may be quantified by self-reporting. Research, clinical practice, the design of systems, and policy-making might all benefit from the use of a psychometric instrument based on this principle.

They propose six Philosophies of Privacy:

- Fatalism

- Moral right
- Nothing to hide
- Something to hide
- Personal responsibility
- Trade-off

«Others have suggested expanding Westin's three clusters of privacy concern into four or five privacy personas.» [165]

Toresson et al. *(2020)* [180] evaluated a privacy impact self-assessment (PISA) app in order to increase awareness of app-related privacy risks. Five personas were created/described (based on Dupree *et al.*'s classification), considering specific privacy vulnerabilities (or threats) in life contexts. Using the privacy impact self-assessment (PISA) app, users could randomly pick an installed app from a database (KAUdroid) of app permission statistics of permissions used to access data on phones, and then chose one of the five vulnerable personas, select partial consent and provide a mitigation action aimed at reducing privacy vulnerabilities. Thus, by PISA app the idea was to increase users' awareness about data sharing and risks while installing apps, using concrete examples of vulnerable personas:

- female e-sport celebrity, known under pseudonym (stalking, sabotage, sponsor loss);
- male well-off elderly citizen with beginning dementia (exploitation, fraud, social exclusion);
- male mid-life professional career, undergoing, cancer treatment (career damage, relationship distress, abusive phone sellers);
- Male, married regional politician with a preference for extramarital affairs (public and private trust engendered, divorce, economic loss);

- teen-age female homosexual in intolerant social environment (discrimination, exclusion, risky contact proposal).

Di Ruscio et al. *(2022)* [59] tried to understand the specific questions useful to differentiate users according to their privacy preferences. An empirical study was conducted using an existing dataset in the fitness field. Specifically, this study focuses on the privacy component of an average user who has no technical awareness of the privacy mechanisms of the digital platforms she smoothly utilizes, but who has a strong moral character and is concerned about privacy. From three to four privacy categories were proposed. While using rigorous settings to secure her data may be beneficial, it may also limit her from fully using all of the software's features and functions. As an alternative, lowering the level of privacy protection may alleviate practical limitations, but at the risk of jeopardizing her personal information. Using the data from the fitness domain study, which was conducted by Sanchez *et al.*. [162] and obtained using a questionnaire and a simulator, the authors took a new approach in this work. They analysed both general and domain-specific questions with the goal of discovering general questions that represent moral attitudes of users and advising privacy choices in accordance with those moral attitudes, among other things. They build and implement a recommender system that presents to users the appropriate privacy suggestions based on their privacy preferences. According to the experimental findings, a compact collection of generic questions performs better than a more complicated domain-dependent set of questions in terms of distinguishing users from each other from a privacy perspective. The suggested recommender system presents users with privacy options, resulting in good prediction accuracy.

They also propose a three/four classification labels to model-driven, top-down selected categories of users:

- *Inattentive*
- *Involved*
- *Attentive*

- *Solicitous*

Where the groups *Involved* and *Attentive* can coincide depending on the clusterization and sample size, analougosly to the *Pragmatists* group proposed by Westin [188].

2.5 Discussion on literature review findings

RQ1.1: *Which study contexts propose privacy categorisations that need to be identified?*

As concerns **RQ1.1**, by this systematic review we showed that there are a relatively few studies that explored the issue of privacy categorization, with a total of **24 papers** emerged from this review. As shown in Figure 2.2 and reported in detail in Table 2.2, there are **9 main contexts/domains** in which the privacy categorization was used or referred in:*General, Economy, Mobile Apps, Health, Social networks, Computing, Location sharing, E-commerce, Internet.* The majority of the studies are focused on general and economical studies. It is important to notice that Westin's works are the majority and are all related in some way to the economic/marketing sector. In early Westin's works the domain/context was predominant in the formulation and the analysis of the privacy categorization, while in more recent works (e.g., Schairer) it emerges that the ethical profile (attitudes, beliefs and in general moral values) influences the performance of the privacy related actions, at least as the domain/context does. It also emerged that privacy attitudes and behaviors may change depending on the level of sensitivity of the domains, such as self-disclosure on social media or GPS-enabled apps versus more sensitive domains such as finance and health. Indeed, according to Nissebaum [138], privacy «...can mean sensitive or intimate information, any information about a person, or only personally identifying information», this depending to the context or domain taken into account. Thus, subjects can be more or less cautious in disclosing personal information according to the domain or to the context in which users behave. For this reason, instead of focusing on one specific domain (e.g., SNS), next studies should adopt a more general-domain approach for better reflecting privacy preferences in terms of personal attitudes rather than of domain-related behaviors. Thus, based on the idea that privacy relies on abstract principles, methods and approaches that capture similarities among human privacy behaviors, regardless of the domain or the context (e.g., health and medicine, and financial), should be investigated [165]. Indeed, privacy is a dimension that reflects ethics and should be driven by the individual's ethical considerations rather than by contextual factors or specific attitudes and practices in a given domain/context [59]. In this vein, it might be desirable to account for stable individual's dispositions (e.g., personality traits) that can be associated with privacy or ethics of privacy in order to categorize users.

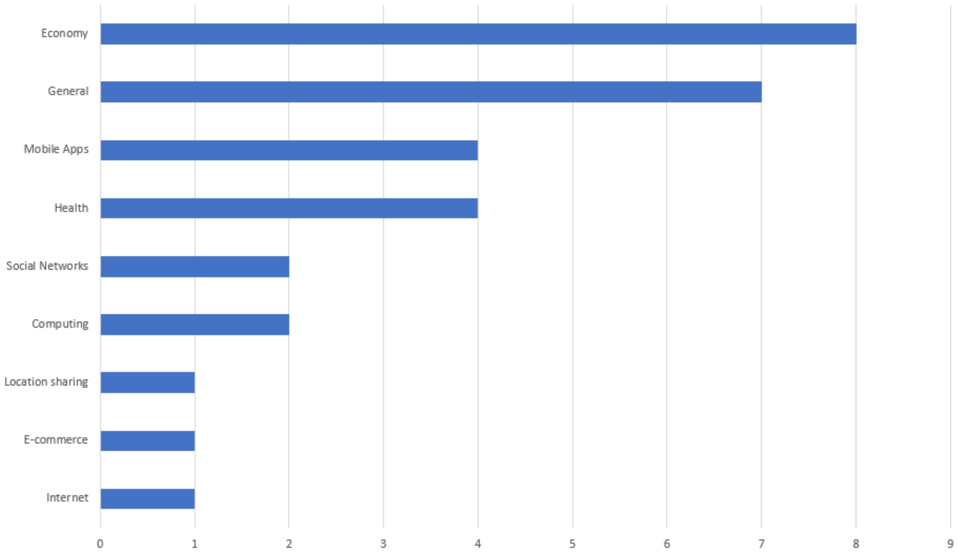


Figure 2.2: Domains/contexts of the studies

RQ1.2: *How can the methodologies and approaches of privacy categorisations be understood?*

Regarding RQ1.2, the literature revealed that a majority of privacy categorizations were derived from various methodologies and tools, such as self-report questionnaires, interviews, and data-set analyses, as illustrated in the Fig. 2.3, although the questionnaire remains the most used tool to measure the privacy behavior. In addition, as shown in Figure 2.4, it also emerged that most of the studies used model-driven approaches, while some used hybrid approaches.

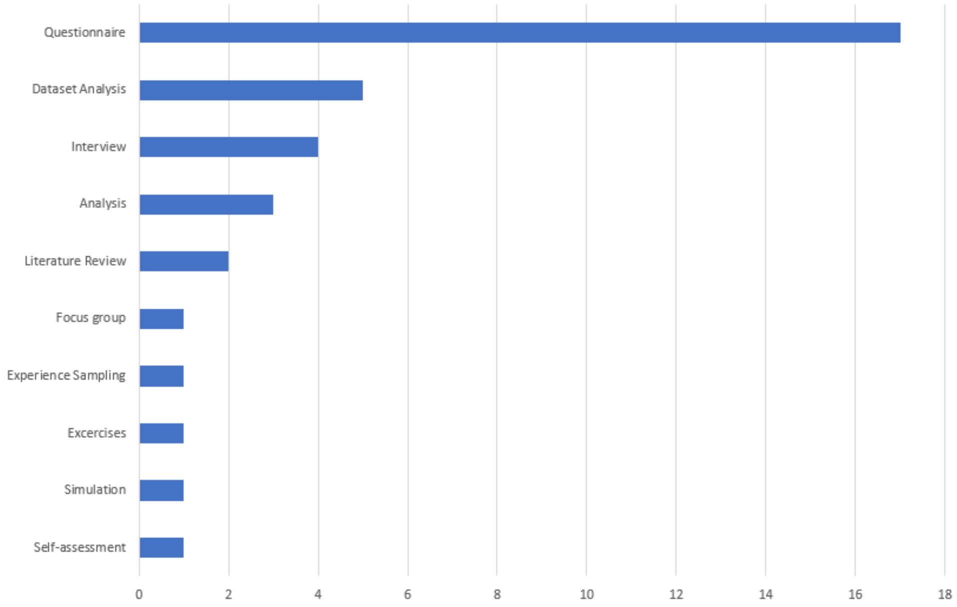


Figure 2.3: Instruments used

This is interesting because the categories personas/philosophies created using hybrid approaches (data- and model-driven e.g., clustering + profiling - and grounded analysis e.g., inductive reasoning) can create a more detailed and specific categorisation than data clusters, modeled segments or parameter-driven profiles. The hybrid approach may balance the limits of the single approaches used. Indeed, on the one hand, when

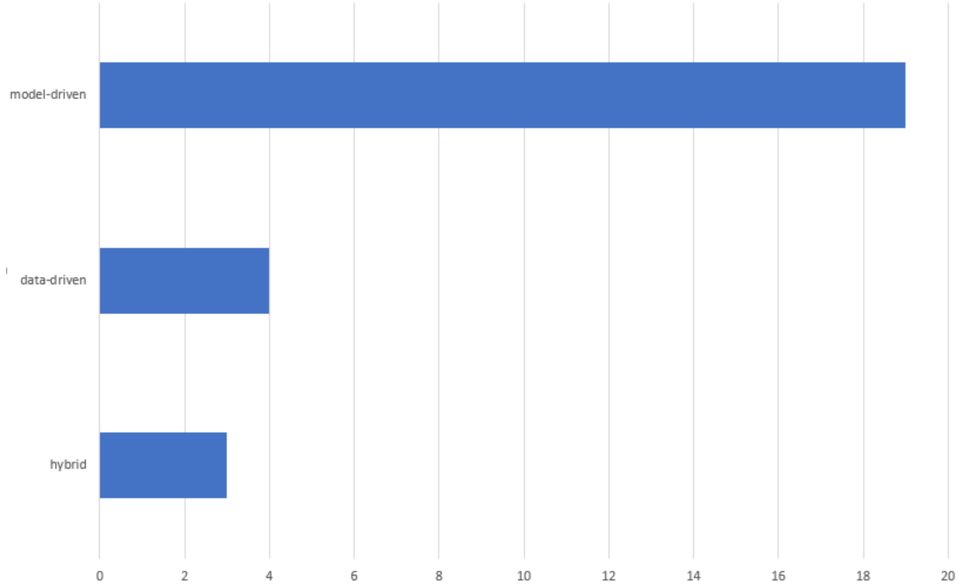


Figure 2.4: Approaches used

using the self-report questionnaires or interviews many of these measures show poor predictive validity of privacy behavior [24, 131], because of the *privacy paradox* phenomenon [140], according to which there is no correspondence between privacy reported attitudes and privacy behaviours. Notably, only Liu *et al.*. [114] used a pure data-driven approach. A totally data-driven approach suffers from the limitations and biases of the analyzed sample, being conditioned by its size [44, 115], homogeneity (e.g., subjects with medical conditions [165]), representativeness (e.g., of a certain country [114]. In this vein, for example, most of the reviewed studies were not balanced in terms of gender [187, 165, 194], whereas the factors of age, education level [60, 194] and technology proficiency for digital behaviours were not appropriately controlled for. To accurately represent the complexity of privacy behavior, it is necessary to refine privacy categories using a wide variety of multidisciplinary approaches and points of view, according to the image. This process of refinement involves the incorporation of insights from various disciplines and methodologies, which can help create a more comprehensive and nuanced understanding of privacy behavior and accommodate the diverse needs and concerns of individuals. The evolution of the analyzed studies suggests that the complexity of the privacy topic cannot be successfully analyzed by applying approaches that take into account only the collection and classification of raw data, especially by those who have a plethora of data at their disposal, such as social networks. In more recent times, the introduction of modern data analysis methodologies, consolidated some categorizations at the expense of others [44, 197]. Based on Dupree [60] and Schairer [165], we can argue that privacy, conceived as a fundamental human right, pervasive in every aspect of human existence and amplified by the enormous possibilities opened up by modern technological means, can only be successfully analyzed and managed through hybrid and interdisciplinary approaches, including also the human science perspective. This means that the interaction between users and the digital world in terms of privacy can be studied combining a computer scientific approach with a more specific human and social scientific approach.

RQ1.3: *How can a map be created to represent the evolution of privacy categorisations?*

Digitization and technology progress has strongly influenced the evolution of the categorization research. We found that the evolution of categorization driven by technology progress has been a defining factor in the studies of the last two decades. The emergence of digital technologies that mediate the interactions between users and societal services in different fields (health, education, e-commerce, etc.) and permit to store, access and analyze large amounts of data, led to more refined solutions to protect users privacy concerns. As a consequence, the process of categorization has become increasingly important in the field of profiling and data analysis. Furthermore, the widespread use of digital technologies has contributed to increase the user's digital fingerprint (in terms of data quantity and variety of sources) thus enabling researchers to identify patterns within their data to make more informed decisions about how to profile and categorize user's preferences and ethics. In addition it

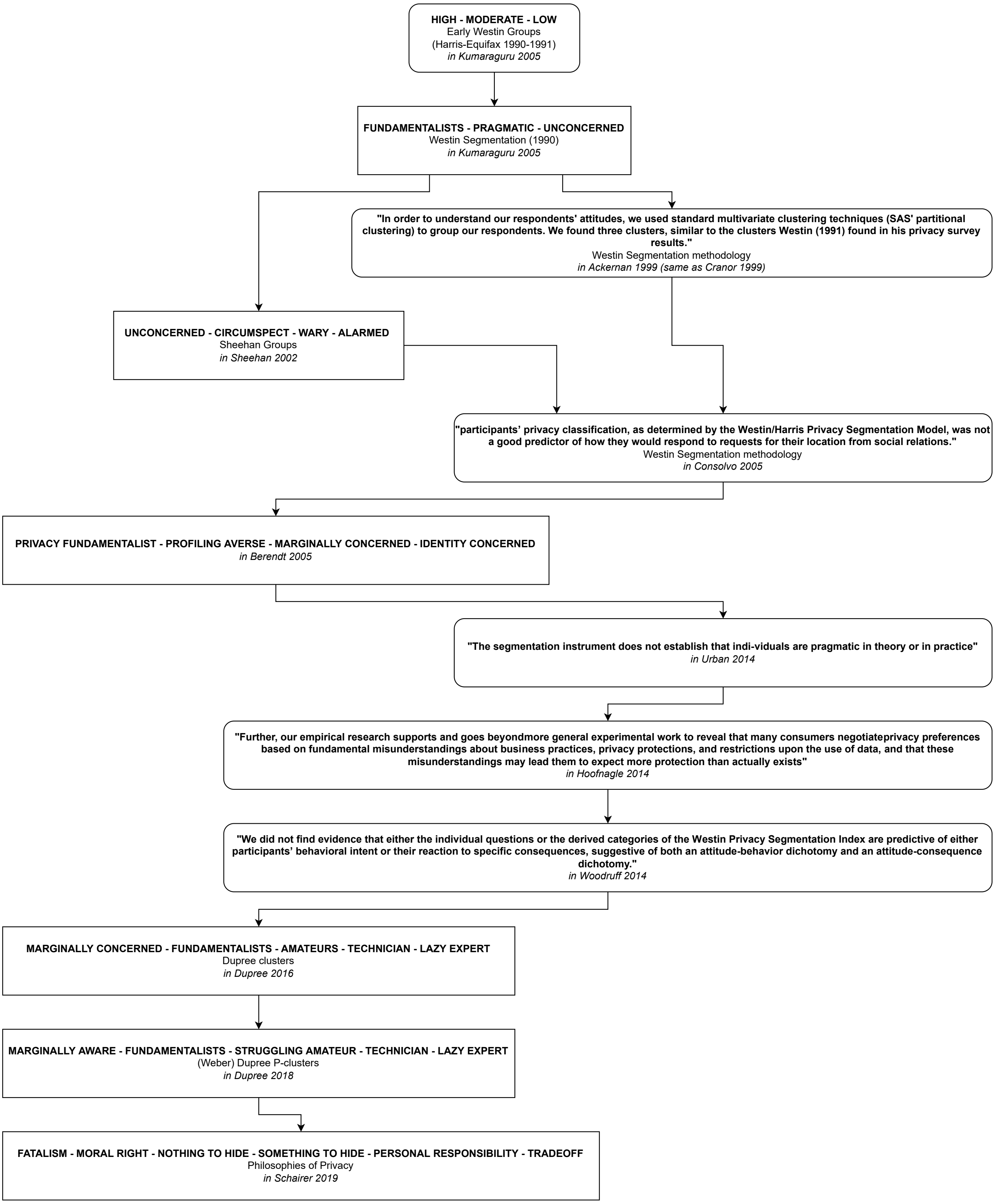

HIGH - MODERATE - LOW
Early Westin Groups
(Harris-Equifax 1990-1991)
in Kumaraguru 2005
FUNDAMENTALISTS - PRAGMATIC - UNCONCERNED
Westin Segmentation (1990)
in Kumaraguru 2005
"In order to understand our respondents' attitudes, we used standard multivariate clustering techniques (SAS' partitional clustering) to group our respondents. We found three clusters, similar to the clusters Westin (1991) found in his privacy survey results."
Westin Segmentation methodology
in Ackernan 1999 (same as Cranor 1999)
UNCONCERNED - CIRCUMSPECT - WARY - ALARMED
Sheehan Groups
in Sheehan 2002
"participants' privacy classification, as determined by the Westin/Harris Privacy Segmentation Model, was not a good predictor of how they would respond to requests for their location from social relations."
Westin Segmentation methodology
in Consolvo 2005
PRIVACY FUNDAMENTALIST - PROFILING AVERSE - MARGINALLY CONCERNED - IDENTITY CONCERNED
in Berendt 2005
"The segmentation instrument does not establish that indi-viduals are pragmatic in theory or in practice"
in Urban 2014
"Further, our empirical research supports and goes beyondmore general experimental work to reveal that many consumers negotiateprivacy preferences based on fundamental misunderstandings about business practices, privacy protections, and restrictions upon the use of data, and that these misunderstandings may lead them to expect more protection than actually exists"
in Hoofnagle 2014
"We did not find evidence that either the individual questions or the derived categories of the Westin Privacy Segmentation Index are predictive of either participants' behavioral intent or their reaction to specific consequences, suggestive of both an attitude-behavior dichotomy and an attitude-consequence dichotomy."
in Woodruff 2014
MARGINALLY CONCERNED - FUNDAMENTALISTS - AMATEURS - TECHNICIAN - LAZY EXPERT
Dupree clusters
in Dupree 2016
MARGINALLY AWARE - FUNDAMENTALISTS - STRUGGLING AMATEUR - TECHNICIAN - LAZY EXPERT
(Weber) Dupree P-clusters
in Dupree 2018
FATALISM - MORAL RIGHT - NOTHING TO HIDE - SOMETHING TO HIDE - PERSONAL RESPONSIBILITY - TRADEOFF
Philosophies of Privacy
in Schairer 2019

Figure 2.5: Segmentations evolution and critics map

has been possible to explore relationships between factors that may have not been previously identified. For example, researchers can now extend and verify the correlations between users' preferences and categories in different and large data sets. As shown in Fig. 2.6, we highlight that in most of the analyzed researches, the profiles also reflect technical behaviours and skills. However, we note that two of the most recent works [165, 59] expand the categorization in ethical and philosophical terms. Thus, although starting from well-defined domains (the medical and the fitness ones, respectively), they propose an ethical profiling from which infer the characteristic traits that can describe the person's behaviour independently from the specific domain.

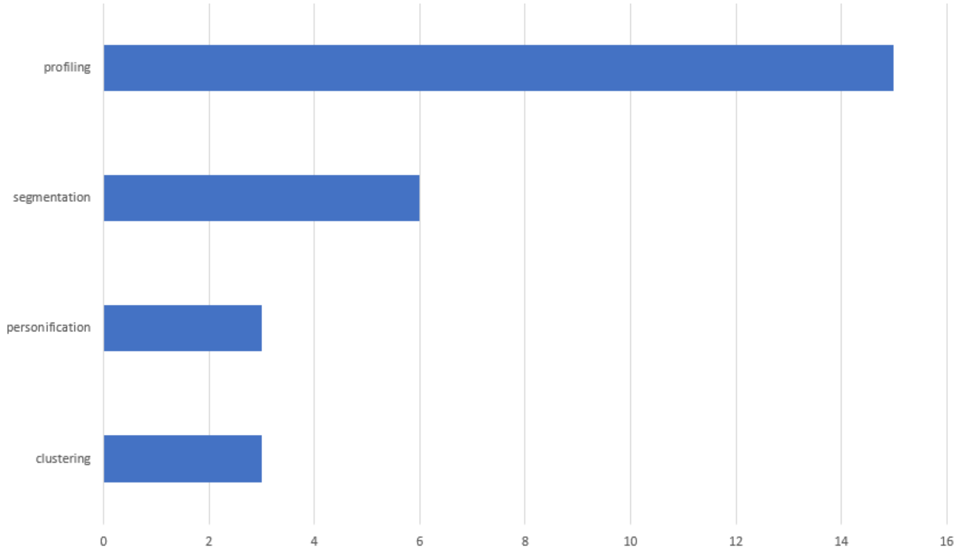


Figure 2.6: Categorizations used

As depicted in Fig. 2.5, starting with the seminal Westin segmentation, which divided the participants into three segments, other studies emphasized the need to define a broader range of profiles, disentangling the *Pragmatic* segment identified by [169, 106] into additional categories. In this vein, although some studies proposed to remove the *Pragmatic* segment and use only two groups, namely *Privacy vulnerable* and *Privacy resilient* [183], other studies claimed the necessity to define digital privacy behavior in more than 3 categories. Using model-driven approaches, some authors identified four clusters, namely *Privacy fundamentalist, Profiling averse, Marginally concerned, Identity concerned* [26], or *privacy conservatives*, *unconcerned*, *fence-sitters*, *advanced users* [111]. Others were able to find five personae using a grounded analysis approach (hybrid): *marginally aware, fundamentalist, struggling amateur, technician, lazy expert* [60, 61]. Then, Wisniewski *et al.*. [194] showed six privacy management categories *privacy maximizers, self-censors, time savers/consumers, privacy balancers, selective sharers, privacy minimalists*, and six privacy proficiency levels *novices, near-novices, mostly novices, some expertise, near-expertise, experts*. Then, Liu *et al.*. [114] identified seven profiles but did not provide a specific label for them. We recall that all the works in this study are related to users' interactions in the digital world, with the exception of *Hoofnagle* et al. *(2014)* [93] which is a theoretical analysis of Westin's work.

2.6 Conclusions on literature review

The issue of the digital privacy categorization has gained an increasing attention in these last years, especially in light of the empowerment and massive diffusion of information technology. In the present systematic review the aim was to investigate the issue of privacy categorization, using 5 keywords, which were elicited by previous studies and references analysis, in order to cover the basic concepts and methodologies used to create the categories.

Categories are constituted by the *elements* investigated through the different **processes**, that is *profile* for **profiling**, *cluster* for **clustering**, *segment* for **segmentation** and *persona* and *philosophy* created by a hybrid data/model-driven and induction process that we will refer to with **personification**. Analyzing the temporal evolution of the definitions of the elements that make up the categorizations, we could see that the terms initially chosen reflected more vague (e.g., Westin [188]) convey a more blurred vision of privacy.

In contrast, in the works published in between years 2000 and 2016, the terms focus mostly on a more technology-oriented view of privacy (e.g, Berendt [26], Dupree [26]). Only in more recent works this state of elaboration is well reflected in the wordcloud (Fig. 2.1), where some specific terms are missing (e.g., ethics/ethical, moral, etc...), suggesting that a better understanding of the role that these concepts can play in categorizing digital users may be profitable.

However, the technological evolution and its diffusion in everyone's everyday life [70] seems to require a more holistic conception of privacy. Although recent regulations (e.g., GDPR [149] - General Data Protection Regulation) requires that any product or service must obtain the user's consent on the way her data will be managed (e.g., by websites and their third parties), the asymmetry in the interaction power between users and digital systems leaves the former unprotected in terms of privacy and security. Bringing the digital technology back to a more ethical dimension that supersedes and surpasses the concept of "online privacy", orienting instead towards a vision of privacy as a fundamental right of human beings appears to be a crucial step in order to preserve the human rights, agency, autonomy and dignity. The latter implies «the recognition of the inherent human state of being worthy of respect», which «must not be violated by 'autonomous' technologies» [146, 15]. To this purpose, approaches that take context into consideration and empowers humans towards autonomous system by exploiting their ethical preferences (Fig. 2.7) like [22] may be pursued.

That would help users to express moral preferences that can also be used to set privacy preferences in the digital world (e.g., disclosing personal information). For this kind of approaches a better understanding of privacy categorizations plays a founding role.

The investigation of digital privacy behaviors, and - by consequence - of categorizations, has theoretical and practical implications. Notably it is possible to:

- relate the new approaches and emerging digital categories to Westin's seminal work.
- understand the complexity of digital privacy behavior, that faces continuous changes and adaptations according to the new regula-

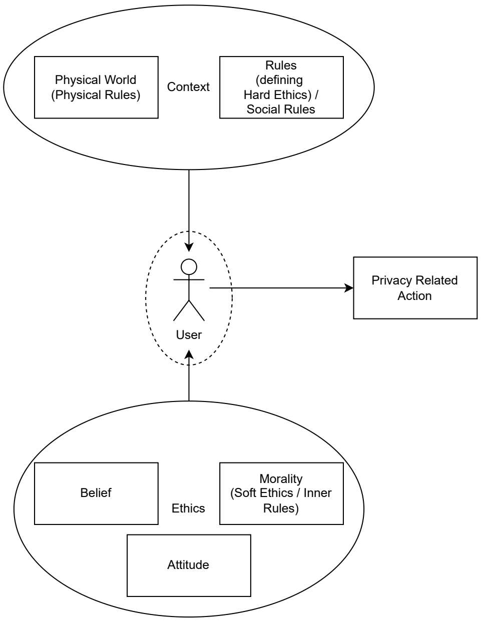

Physical World (Physical Rules)
Context
Rules (defining Hard Ethics) / Social Rules
User
Privacy Related Action
Belief
Ethics
Morality (Soft Ethics / Inner Rules)
Attitude

Figure 2.7: Ethics and context influence on the user while performing an ethical sensitive action

tions, opportunities, digital skills and societal awareness.

- adapt the digital privacy protocols and new technologies to categories in order to satisfy the new challenges that the society requires. This implies a better designing of digital technologies, according to the users' digital privacy preferences.

- encourage digital privacy education and nudging, improving users' knowledge and proficiency as well as making easier to carry on a specific behavior.

Future works in privacy profiling and categorization should be aimed to define privacy categories using a variety of multi-disciplinary approaches and perspectives, pursuing more predictive categories, such as *personae* and *philosophies*. This would better guarantee a correspondence between users' general privacy beliefs and users' behavior when expressing privacy preferences.

Chapter 3

Eliciting user's privacy profile: mobile apps domain data set analysis

Concerns about the ethics and the people's right to privacy are at the forefront of the issues brought up by our increasingly digital culture. Users are routinely profiled by software programs, which results in the need for users to effectively control their privacy settings. This requirement is also enforced by regulations. Protecting individually identifiable information and expressing personal opinions for ethical issues both require users to effectively handle the privacy settings of software. Artificial intelligence technologies that allow people to engage with the digital world in a way that reflects their own personal ethical values have the potential to be significant boosters of a trustworthy digital society. We concentrate on the privacy aspect, and via an empirical investigation conducted on an already existing dataset gathered from the fitness area, we offer a step in the direction outlined above. Our goal is to determine which group of inquiries is most suited to classify consumers according to the level of privacy protection they need. According to the findings, a condensed collection of semantically driven questions (relating to domain-independent privacy preferences) is superior than a complicated domain-dependent one for helping to differentiate people. This provides support for the premise of the research, which states that the most important piece of data to gather is participants' moral opinions. In light of the findings, we will create a recommender system in order to provide consumers with appropriate suggestions in regard to their privacy-related options. After that, we demonstrate how the suggested recommender system may offer users with suitable settings to achieve a high level of accuracy.

3.1 Preface

Privacy and ethics of citizens are at the core of the concerns raised by our increasingly digital society. Profiling users is standard practice for software applications triggering the need for users, also enforced by laws, to properly manage privacy settings and moral preferences. This deals with the way users give their consent to storing, sharing to third parties, as well as disseminating sensitive personal information and express moral preferences like, for example, ticking to pay a *decarbonization tax*.

In this chapter, we focus on the privacy dimension of an ordinary user with little technical knowledge of the privacy mechanisms of the digital systems she seamlessly uses but with an evident moral character. While choosing strict settings may help protect her data, this may prevent the complete availability of the functionalities provided by the software. In contrast, loosening privacy settings mitigates the restriction on functionalities, but it may come with the price of compromising her data privacy. In this respect, Artificial Intelligence (AI) technologies can empower the user in maintaining a reasonable trade-off between accessibility and protection, and reflecting the user privacy preferences can be the key enabler of a trustworthy digital society. Understanding the commonalities and differences among users based on profiles has been among the main issues in data privacy research [188, 106, 200]. Categorizing profiles contributes to better identification of users' behaviors and supports administrators in comprehending privacy choices. At the same time, personal profiles may enable the design of functionalities that help users set privacy preferences of the digital technologies they use. Various proposals to categorize or group end-users into clusters based on their security or privacy attitudes/behaviors in specific domains have been made [156, 108]. Users' preferences were analyzed in an extensive study [111] on permission settings from real Android mobile users to recommend personalized default settings. Sanchez *et al.* [162] analyzed user-privacy preferences in the fitness domain employing a specifically designed questionnaire consisting of both domain-specific and general questions to recommend personalized privacy settings for the fitness apps.

Though a lot of achievements have been reported, as discussed in [113], we believe that there is still the need to understand how to characterize user's privacy behavior in a general setting. Indeed privacy is a dimension of ethics and should be part of the ethical profile of a user and driven by ethical consideration rather than by contextual attitudes or practices in given domains. For example, relying on the analysis of current or past users' preference settings as in [111] does not guarantee a correspondence between what users believe as their general privacy profile and what they actually (can) do when setting privacy preferences. Moreover, data privacy awareness in the digital society is only recently exiting the specialists" fields (legal, ethical, economic, social) to impact the wider society. The pandemic has also dramatically advanced the penetration of digital technologies in the society from market to education [126, 144]. This means that a large body of collected data on privacy settings may not reflect the attitude and attention to privacy that present users have and will have in the future.

In this work, we explore a different research direction by relying on the data of the study in the fitness domain [162] that were collected by means of a questionnaire and a simulator.[1] We analyze both general and domain-specific questions with the aim of *(i)* identifying general questions that reflect moral attitudes of the users;

and *(ii)* recommending privacy preferences accordingly. Moreover, we design and implement a recommender system [154] to provide users with suitable recommendations with respect to privacy choices. The experimental results are positively interesting, revealing that a compact set

[1] We thank Prof. Dr. Ilaria Torre, University of Genoa (Italy) for providing us with the privacy dataset [162].

of general questions helps distinguish users better than a more complex domain-dependent one. We also show that the proposed recommender system provides relevant settings to users, obtaining high prediction accuracy.

The main contributions of this chapter are summarized as follows.

- We investigate which sets of (general) privacy questions are more relevant for classifying users with respect to their privacy moral preferences.
- By means of an empirical evaluation, we show that self-assessment about privacy attitudes given by users does not reflect the way they act in practice.
- We develop PisaRec, a recommender system to provide suitable privacy settings that reflect user preferences. This aims to help users relieve the burden of setting privacy configurations when they go online.

We organize the chapter into the following sections. In Section 3.2, we present a motivating example and a categorization of privacy profiles. Section 3.3 describes the proposed approach which makes use of both unsupervised and supervised learning to handle user profiles. The methods used to evaluate our approach are detailed in Section 3.4. We report and analyze the experimental results in Section 3.4.4. Discussion related to the limitations and threats to validity are provided in Section 3.5. We review related work in Section 3.6. Finally, Section 6 sketches future work and concludes the chapter.

3.2 Background

The following example illustrates the need for personalized automated privacy assistance that a user interacting with multiple systems at a time may require. Then we briefly report the most relevant aspects for our research taxonomies for privacy profiles proposed in the literature.

3.2.1 Motivating example

After a long day at work, Alice is at the subway station. After the pandemic outbreak, she will meet pals at the cinema. She is on time but learns that she cannot buy a ticket from the subway station attendant due to rigorous hygiene regulations. In addition, vending machines are out of commission for contact-less technology upgrades. Instead, a QR code and simple instructions to buy an electronic ticket online are posted in front of the vending machines.

Her train is about to arrive, she opens her camera app and frames the QR code. The site structure appears in a split second, but as Alice scrolls down to find the ticket she needs, a popup asks for her privacy settings. Above a very long list of radial button options about disclosing GPS position, information about her mobile phone, consent to save various types of cookies on her device, share her list of contacts, etc., she is presented with three buttons: *accept all*, *strictly necessary*, *decline all.*

Alice is very concerned about her privacy, and when not strictly necessary for the purpose she wants to perform, she does not wish to disclose private information. Since the service she is asking for is simple as asking for a one-ride ticket, she clicks *decline all.* The next page seems to load slowly, images and structure are shown in a non-adaptive way, so she has to pinch-in to zoom and scroll to read the text that informs her that a cryptographic key used for her session management cannot be stored due to her preferences, so the session is not secure also the page asks her to choose language, timezone, type of device and the web browser she is using, payment options, etc. While reading, Alice realizes that her train is about to arrive at the station. So, she decides to click the back button on her browser, reload the page and click *strictly necessary* when prompted. The site then stays fast and steady, adapted to the display of her device, prompting if she wants to take a one-ride ticket or a full day one. Her mobile wallet handles the payment instantly, and she receives her ticket just before the train comes. On time to the cinema, Alice enjoys the film with her friends, soon forgetting the online ticket purchase experience. Her preferences are saved on her phone, so she will buy train tickets quickly and easily in the future.

Table 3.1: Privacy categories according to different taxonomies (Listed in chronological order).

How important is privacy to you?	NOTHING	LITTLE		QUITE		VERY
Segmentation [188	Unconcerned	Pragmatists				Fundamentalist
Privacy Personas [60]	Marginally concerned	Amateurs		Technicians	Lazy Experts	Fundamentalists
Philosophies [165]	Fatalism	Nothing to hide	Something to hide	Trade-off	Personal Resp.	Moral right
Privacy Clustering [162]	Unconcerned	Socially active		Health-focused	Minimal	Anonymous (Strict)
Self-assessment [162]	Conservative	Unconcerned		Fence-Sitter		Advanced Users
Our proposed categorization	INATTENTIVE	INVOLVED/ATTENTIVE				SOLICITOUS

Alice does not know that the *strictly necessary* option, although excluding third-party tracking and marketing, includes all alternatives that are strictly essential to all services offered by the booking site, including – the lower price inter-city ticket that requires GPS tracking, the discounted price for kids that requires age disclosure, discount for army and state officials who must check other installed mobile applications, as well as the train pass app to see if the ticket is part of a booklet, etc.

Analogously to various studies notably Liu *et al.* [114], we believe that a software technology should assist Alice in automatically selecting the options that, on the one hand, are needed for what she wants to do and, on the other hand, are compliant with her moral preferences.

In this work, we show that it is possible to protect users by first understanding their privacy profiles, which can be automatically identified by considering a small set of general and domain-independent questions that are shown to be enough to reflect the user's moral attitude. Thus, our approach is to categorize personal privacy profiles from an ethical perspective [22]. Profiles can then be used to automate app and web settings, leveraging recommender systems like in this chapter or other technologies.

3.2.2 Categorizations of privacy profiles

Table 3.1 gives a summary of the most notable taxonomies of privacy categories. Starting from the question: "*How important is privacy to you?*" from left to right of the table, an increasing level of privacy concerns is shown.

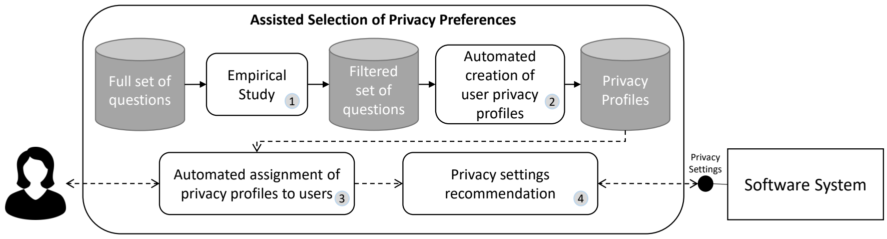

Assisted Selection of Privacy Preferences
Full set of questions
Empirical Study 1
Filtered set of questions
Automated creation of user privacy profiles 2
Privacy Profiles
Automated assignment of privacy profiles to users 3
Privacy settings recommendation 4
Privacy Settings
Software System

Figure 3.1: The proposed approach.

Westin [188] proposed the first categorization of user profiles with three levels, i.e., *Unconcerned*, *Pragmatists*, and *Fundamentalist*. Since then, there have been other studies that follow up and develop this initial taxonomy. In particular, Dupre *et al.* [60] expanded it proposing five categories: *Marginally concerned*, *Amateurs*, *Technicians*, *Lazy Experts*, and *Moral right*. Schairer *et al.* [165] came up with even more, i.e., six categories, where the answer *Little* is split into *Nothing to hide*, and *Something to hide*; and *Quite* is made of *Trade-off*, and *Personal Resp.* Recently, Sanchez *et al.* [162] proposed a more compact categorization, where users are grouped into four categories, *Privacy concerns*, *Unconcerned*, *Fence-Sitter*, and *Advanced Users*. As it appears from the table, category refinement happens in the middle category and may depend on the application domain as well as on the amount of input data. Further categories are contextual and can be obtained as specialization of personal profiles based on the single user's experience. Therefore, as starting point three categories provide three clearly distinguishable moral attitudes.

The categorization we propose names the three clusters as *Inattentive*, *Attentive*, and *Solicitous*. While *Inattentive* means that users do not care about privacy, *Solicitous* corresponds to an opposite attitude, where users are completely aware of privacy issues. The *Attentive* category is something in between, and covers both *Little* and *Quite* answers to the question "*How important is privacy to you?*"

3.3 Proposed approach

Typically, users specify privacy preferences by directly interacting with the privacy settings provided by the used software. Similar to other techniques [22, 114, 113] we propose an approach that relies on a software layer that automatically identifies privacy profiles and interacts with the user or the software system to recommend privacy preferences accordingly.

Concerning what we present in this chapter, the assisted selection phase of privacy preferences started on training data consisting of general, domain-specific, and app-specific answers given to the questions defined in [162] (see *"Full set of questions"* in Figure 3.1). We have empirically analyzed the full set of questions to identify the corresponding subset (consisting of general questions) that is sufficient to automatically identify our three user privacy profile categories, i.e., *Inattentive*, *Attentive*, and *Solicitous* (see activity ① in Figure 3.1).

The automated creation of user privacy profiles phase (see activity ② in Figure 3.1) relies on an unsupervised clustering module, which can automatically group users in the training data. The automated assignment of privacy profiles to users phase (see ③) relies on a supervised classifier using a feed-forward neural network to automatically assign to the given user the corresponding privacy profile among one of those identified in ②. Finally, a recommender system is used to further validate the activities ① and ②, and to provide users with privacy settings recommendations (see activity ④) according to the privacy settings of other users belonging to privacy profiles as detected in ③.

Details on the results obtained from the performed `Empirical Study` are given in the *Experimental results* section, whereas the activities ②, ③, and ④ are described as follows.

3.3.1 Automated creation of user privacy profiles

To automatically create user privacy profiles, we employed a clustering process by relying on the graph-based representation of users and privacy settings as shown in Figure 3.2.

This representation is also used by the developed neural network for classifying users presented in Section 3.3.2. The graph represents the relationship between users and settings. A directed edge is formed between a user and a setting if the user already turned on the setting.

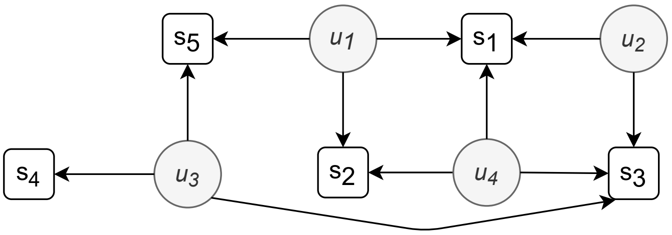


Figure 3.2: Graph representation of users and privacy settings.

Each user u is represented by a vector $\phi = (\phi_1, \phi_2, .., \phi_F)$, where ϕ_i is the weight of term s_i, computed as the *term-frequency inverse document frequency* value as follows:

$$\phi_i = f_{s_i} \times log(\frac{|P|}{a_{s_i}}) \tag{3.1}$$

where f_{s_i} is the number of occurrence of s_i with respect to u, it can be either 0 and 1 since there is a maximum of one s_i connected to u by the corresponding edge; $|P|$ is the total number of considered users; a_{s_i} is the number of projects connecting to s_i via an edge.

The similarity between two users u and v is computed using their corresponding feature vectors $\phi = (\phi_1, \phi_2, .., \phi_F)$ and $\omega_=(\omega_1, \omega_2, .., \omega_F)$

$$sim(u, v) = \frac{\sum_{t=1}^{n} \phi_t \times \omega_t}{\sqrt{\sum_{t=1}^{n} (\phi_t)^2} \times \sqrt{\sum_{t=1}^{n} (\omega_t)^2}} \tag{3.2}$$

where n is the cardinality of all settings that were set to 1 by both u and v. Intuitively, u and v are characterized by using vectors in an n-dimensional space, and Equation 3.2 measures the cosine of the angle between them. As an example, in Fig. 3.2, we see that the two users u_2 and u_4 are similar since they both set two settings s_1 and s_3.

A set of n users is grouped into κ pre-defined number of clusters, with the aim of maximizing both the similarity among instances within a single cluster, and the dissimilarity among independent clusters. To this end, we calculate the distance between every pair of users and feed as the input for the clustering engine. The K-medoids algorithm [148] has been chosen to group users into clusters due to its simplicity and efficiency.

In the clustering process, the distance scores, computed as $d_C(u, v) = 1 - sim_C(u, v)$, are used to assign users to clusters. Initially, a set of medoids (users) is generated randomly, then a medoid is selected as the user in the cluster that has minimum average distance to all the other users in the cluster. Afterwards, users are assigned to the cluster with the closest medoid, using a greedy strategy [148].

3.3.2 Automated assignment of privacy profiles to users

Supervised learning algorithms can simulate humans' learning activities, mining knowledge from labeled data and performing predictions for unknown data [86]. Among others, neural networks have been widely adopted in various applications, including pattern recognitions [28], or forecasting [199]. A feed-forward neural network consists of connected layers of neurons, where the output of a layer is transferred to the next layer's neurons, except for the output layer.

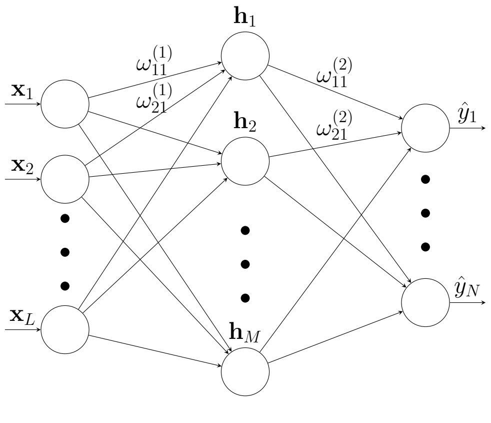


Figure 3.3: A three-layer neural network.

We built a feed-forward neural network to classify users into different privacy groups, using preferences as features. The network consists of three layers explained as shown in Figure 3.3. The input layer has L neurons, being equal to the number of input settings, i.e., $X = (x_1, x_2, ..., x_L)$. The middle layer consists of M neurons, i.e., $H = (h_1, h_2, ..., h_M)$, M can be configured during the evaluation. There are κ neurons in the output layer, corresponding to κ output categories, i.e., $\hat{y} = (\hat{y}_1, \hat{y}_2, .., \hat{y}_\kappa)$. The predicted value $\hat{y}_k$ for neuron k of the output layer is computed to minimize the error between the real values and the predicted ones. As discussed in the *Experimental results* section, the conceived neural network has played an important role in the performed analysis, especially to understand to what extent self-declared privacy profiles reflect the actual user category.

3.3.3 Privacy settings recommendation

We conceptualize PisaRec, a **Pri**vacy **s**ettings **a**ssistant running on top of a **Rec**ommender system to provide users with suitable data protection configurations. PisaRec works based on the assumption that *"if users of the same privacy profile already share some common privacy settings, then they are supposed to share additional similar settings"* [164]. In this way, we utilize the proposed graph-based representation to model the relationship among users and use a collaborative-filtering algorithm [12] to recommend missing settings. To feed as input for the recommendation engine, we adopt the *user-item* paradigm [139], in which each user corresponds to one row, a column represents each setting. In this way, a cell in the matrix dictates the rating given by a user to a setting. The two values 0 and 1 correspond to *deny* and *allow*, respectively. An example of a user-setting matrix for the set of four users and five settings is as follows: $u_1 \ni s_1, s_2$; $u_2 \ni s_1, s_3$; $u_3 \ni s_1, s_3, s_4, s_5$; $u_4 \ni s_1, s_2, s_4, s_5$. Accordingly, the user-item ratings matrix built to model the occurrence of the settings is depicted in Figure 3.4.

The following collaborative-filtering formula is utilized [164] to predict the inclusion of a setting s_i for user u:

$$r_{u,s_i} = \overline{r_u} + \frac{\sum_{v \in topsim(u)} (r_{u,s_i} - \overline{r_v}) \cdot sim(u,v)}{\sum_{v \in topsim(u)} sim(u,v)} \tag{3.3}$$

$$\begin{array}{cc} & \begin{array}{ccccc} \mathbf{s}_1 & \mathbf{s}_2 & \mathbf{s}_3 & \mathbf{s}_4 & \mathbf{s}_5 \end{array} \\ \begin{array}{c} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \\ \mathbf{u}_4 \end{array} & \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

Figure 3.4: A user-setting matrix.

where $\overline{r_u}$ and $\overline{r_v}$ are the mean of the ratings of u and v, respectively; v belongs to the set of *top-k* most similar users to u or neighbour users, i.e., $topsim(u)$; $sim(u, v)$ is the similarity between u and a similar user v, computed using Equation 3.2.

The clusters obtained from the previous section allow us to identify users with similar privacy preferences. Based on the obtained categorization, given an input user, the neural network assigns the user to a specific category. Afterward, we build a graph only for this category following the paradigm in Figure 3.2. Such a sub-graph contains fewer nodes and edges than a full graph for all categories, aiming to optimize the computation. On top of this, PisaRec recommends missing settings to users. The outcome of the computation is a ranked list of probable settings, and we select the top-N of them to present as the final recommendations.

3.4 Evaluation

To study the proposed approach's performance, we first introduce three research questions. Afterward, we describe the dataset and metrics used in our evaluation.

3.4.1 Description of the research questions

The following research questions are considered to evaluate our proposed approach.

RQ2.1: *How well does the users' self-assessment reflect their privacy category?*

As users in the considered dataset [161] have been allowed to self-assess their privacy category, we examine if such a self-evaluation reflects their real category.

RQ2.2: *Which sets of questions are relevant for assessing privacy concerns?*

We are interested in finding the set of questions that can better distinguish between user profiles. For this research question, we cluster the users with different sets of features, and identify the one that brings the best clustering solution. The aim is to find a set of privacy questions that better represents the user profiles.

RQ2.3: *To what extent is a recommender system able to utilize the obtained categorization in recommending relevant privacy settings to users?*

We investigate how well the conceived recommender system learns from existing profiles, providing users with additional configurations that reflect their preferences.

3.4.2 Dataset

We opted for an existing dataset that has been collected through a domain-specific survey about the usage of a fitness app including user privacy preferences [162].

As shown in Table 3.2, there are 444 data entries (questions and other type of data, like timing) which have been divided into three main groups as follows:

Table 3.2: Summary of the dataset.

Questions/Data	**Alias**	**Description**	**# entries**
Domain specific	D	Questions related to the specific domain (Fitness)	**202**
D Subset 1	DP1	Subset of the D set consisting of privacy relevant questions	123
App related	A	Questions related to the mobile application and the specific software context	**113**
A Subset 1	AP1	Subset of the A set consisting of privacy relevant questions	65
A Subset 2	AP2	Subset of the A set that includes only generalizable questions	6
D + A Subset 0	S0	Privacy related questions from the D and the A sets (DP1+AP1)	188
Generic	G	Generic questions not specifically related to the domain (fitness) or the application/-software context (mobile app)	**129**
G Privacy Subset 1	GP1	Subset of the G set consisting of privacy relevant questions	110
G Subset 1	G1	Subset of the G set consisting of questions related to the disclosure of information about user's identity with the app	35
G Data 2	G2	Data concerning the time spent by the users to answer the questionnaire, play with the simulator, and the sum of the two	3
G Subset 3	G3	Subset of the G set consisting of questions related to the user's identity	19
G Subset 4	G4	Subset of the G set consisting of questions related to the disclosure of private information with the app	56
G Subset 5	G5	Subset of the G set consisting of questions related to the concerns about privacy	16
Full dataset	DATA	Data collected with the questionnaire and the simulator (D+A+G)	**444**

- *Domain specific*: This is the set of questions being explicitly related to the fitness activity. There are a total of 202 questions in this category.

- *App related*: These questions are about the use or setting of the app, consisting of 113 questions.

- *Generic*: This set of questions consists of generic questions that are not related to other groups. There are 129 generic questions in total.

3.4.3 Evaluation metrics

- **Compactness**. The metric measures how closely relevant the users within a cluster are [117]. In this respect, a lower value represents a better clustering solution and vice versa.

- **Silhouette**. It measures how similar a user u is to all the remaining users of the same cluster [117], computed using the following formula:

  $$s(u) = \frac{(b(u) - a(u))}{max\{a(u), b(u)\}} \tag{3.4}$$

  where $a(u)$ is the mean distance between u and the others, $b(u)$ is the minimum mean distance. A silhouette value falls into the range [-1,..+1], where a higher score means a better clustering solution.

  Furthermore, we also use Precision, Recall, ROC curve and AUC to study the performance of the proposed approach.

  First, there are the following definitions: True positive (TP) is the settings that match with ground-truth data; False positive (FP) is the recommended settings but do not match with the ground-truth data; False negative (FN): the settings that should be recommended, but they are excluded. Then, the metrics are as follows:

- **Precision and Recall**. Precision measures the fraction of the number of settings properly classified to the total number of recommended items and Recall (or true positive rate – TPR) is the ratio of the number of correctly classified items to the total number of items in the ground-truth data. The metrics are defined as follows:

  $$P = \frac{TP}{TP + FP} \tag{3.5}$$

  $$R = \frac{TP}{TP + FN} = TPR \tag{3.6}$$

- **False positive rate (FPR)**. This metric measures the ratio of the number of items that are falsely classified into a category **c**, to the total number of items that are either correctly not classified, or falsely classified into the category:

  $$FPR = \frac{FP}{TN + FP} \tag{3.7}$$

- **ROC curve and AUC**. The relationship between FPR and TPR

Figure 3.5: ROC curves with generic questions.

is sketched in a 2D space, using a receiver operating characteristic (ROC) [69], which spans from (0,0) to (1,1). An ROC close to the upper left corner represents a better prediction performance.

3.4.4 Experimental results

This section reports and analyzes the experimental results by answering the research questions introduced this chapter.

RQ2.1: *How well does the users' self-assessment reflect their privacy category?*

In the dataset [162] considered in our evaluation, each user has assigned themselves to one of the following four groups: *Privacy Conservative* (Class 0), *Unconcerned* (Class 1), *Fence-Sitter* (Class 2), and *Advanced User* (Class 3).

We investigate if the self-assessment is consistent, i.e., if all the users properly perceive their real privacy category. This is important since a proper self-clustering can be utilized in additional profiling activities.

We conducted evaluation using the conceived neural network as the classifier. Such a technique has been successfully applied to classify various types of data, e.g., text [129], chemical patterns [31], metamodels [133], to name a few. Similarly, we use the privacy settings as features, and the labels specified by humans to train the classifier. We opt for the ten-fold cross validation technique [103], where the dataset is split into ten equal parts, and the evaluation is done in ten rounds.

In this way, the neural network attempts to learn from the data labelled by users and we investigate if there is a concrete pattern in the considered data.

The evaluation metrics are computed on the test set, i.e., for each user the network predicts a label, which is then compared with the self-assessed label to evaluate the performance. Finally, ROC curves are sketched by combining the scores obtained from all the ten folds.

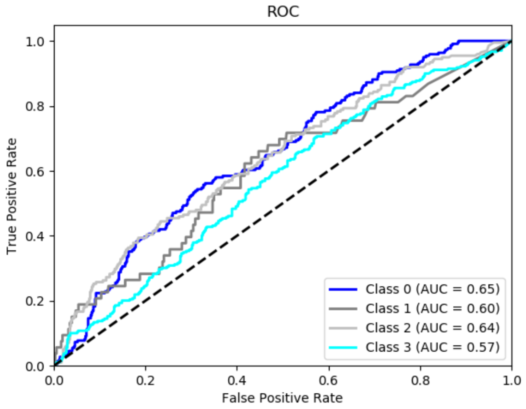


Figure 3.6: ROC curves with domain specific questions.

Figure 3.5 and Figure 3.6 depict the ROC curves obtained from the classification results for generic and domain specific questions. It is evident that the classifier achieves very low prediction performance on both configurations. In particular, the curves bend over the diagonal line, being close to a random guess. Moreover, the AUC values of the four categories are always lower than 0.65. In other words, we encounter *negative results*, where the neural network fails to predict a proper category for a user. These results suggest that there are noises in the training data [201], which could possibly be both in the features and the labels.

To confirm the hypothesis, we measure the similarity between each user and all the remaining others. Interestingly, we found out that 96.20% of the users have very similar users in completely different self-assessed categories.

This demonstrates that while users share similar preferences, they classify themselves differently, causing a low prediction performance for the neural network.

Answer to RQ$_{2.1}$. The self-assessment given by users does not reflect their real privacy category: Users with highly similar settings perceive themselves as completely different groups. In practice, this means administrators should not rely on such a self-categorization, but have to perform privacy profiling on their own.

RQ2.2: *Which sets of questions are relevant for assessing privacy concerns?*

As seen in **RQ**2.1, the self-assessment given by users is not consistent, thus it is necessary to find another way to group users into clusters. We performed experiments on different subsets of the questionnaire to study the influence of each set on the clustering results. The ultimate aim is to identify a set of questions that helps classify users better.

In particular, we are interested in analyzing the following groups of questions:

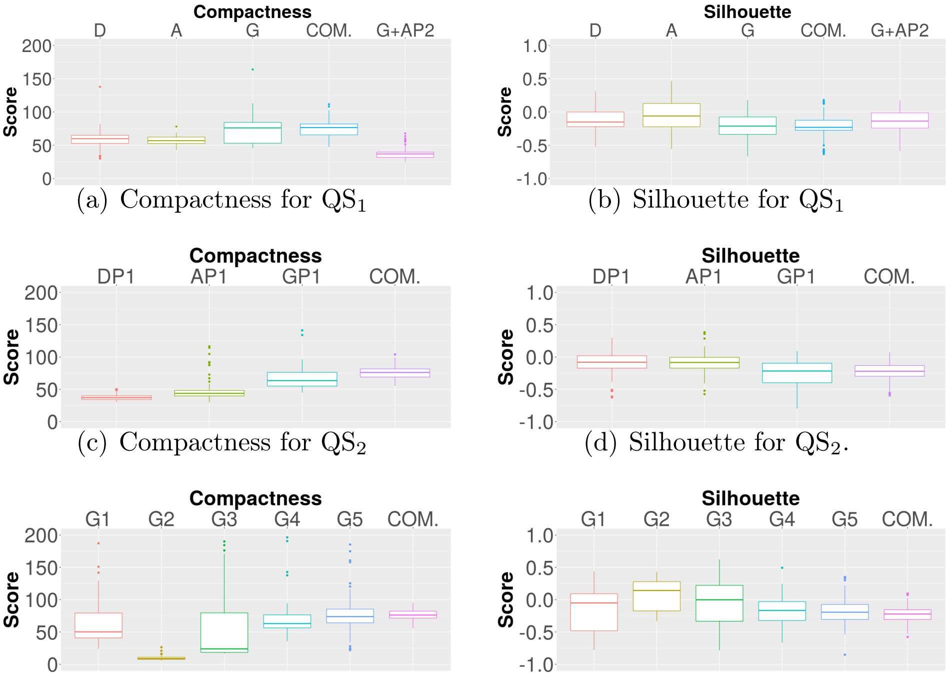
- QS_1: It is a set of question sets as follows: Domain specific (**D**); App related (**A**); Generic (**G**) and their combination (i.e., **D**+**A**+**G**) named **COM.**. Furthermore, we also include the set **G+AP2** where **AP2** contains generalizable questions like

  "Do you believe the company providing this fitness tracker is trustworthy in handling your information?" Indeed, this is to ask if the company is trustworthy, and therefore we consider it as a general question.

  QS_1 permits to compare compactness and silhouette performances between a single set and combinations of all questions.

- QS_2: It is a set of questions sets as follows: **DP1**, **AP1**, and **GP1** that are the subsets of **D**, **A**, and **G** consisting only of privacy relevant questions, respectively. **COM.** is the union of the three subsets, i.e., **COM.** = **DP1+AP1+GP1**. QS_2 permits to understand the actual influence of the privacy-related questions.

- QS_3: It consists of subsets of generic questions **G** defined as follows: **G1** are the questions related to disclosure of information about user's identity with the app; **G2** the questions related to the time spent by the user in completing the survey; **G3** the questions related to user's identity; **G4** the questions related to disclosure of private information with the app; **G5** the questions related to concerns about privacy. **COM.** is the combination of all the subsets,



(a) Compactness for QS_1

(b) Silhouette for QS_1

(c) Compactness for QS_2

(d) Silhouette for QS_2.

(e) Compactness for QS_3

(f) Silhouette for QS_3.

Figure 3.7: Compactness and silhouette scores.

i.e., **COM.** = **G1**+**G2**+**G3**+**G4**+**G5**. QS_3 is to ascertain the influence of the generic questions with respect to the overall set of questions.

We compute and report for each set the corresponding compactness and silhouette scores. Figure 3.7(a), Figure 3.7(c), and Figure 3.7(e) report the compactness scores computed for the three question sets.

As it can be seen in Figure 3.7(a), using **A** as input yields the most compact clusters. In particular, most of the scores are smaller than 40. When domain specific questions (**D**) are used as the features, we also obtain low compactness scores, albeit being larger than using **A**. If only generic questions, i.e., **G**, are utilized, worse clustering solutions are seen. When comparing the results obtained by using **G** with those of using **G+AP2**, we can see that adding **AP2** to **G** contributes to a better clustering. Concerning QS_2 where only privacy relevant questions are considered, we see that using domain specific privacy relevant questions (**DP1**) allows us to gain the most discriminative clusters. Using the subset of privacy relevant questions, i.e., **AP1** is also beneficial to the clustering of user profiles.

For QS_3, there are comparable clustering solutions when using the features sets $\mathbf{G}_1$, $\mathbf{G}_3$, $\mathbf{G}_4$, and $\mathbf{G}_5$. The best clustering is obtained with $\mathbf{G}_2$.

The silhouette scores in Figure 3.7(b), Figure 3.7(d), and Figure 3.7(f) further enforce the compactness ones. **A** is the feature set that achieves the best silhouette for QS_1. Adding **AP2** to **G** helps achieve a better clustering solution, compared to using only **G**.

Answer to $\mathbf{RQ}_{2.2}$. According to the performed evaluation, generic questions plus generalizable ones (i.e., **G+AP2**) provide the best clustering solution.

RQ2.3: *To what extent is a recommender system able to utilize the obtained categorization in recommending relevant privacy settings to users?*

An issue with clustering is whenever there is a new user to be classified, it is necessary to re-run the whole process. This is a time consuming phase, especially where there is a large number of users. Thus, we propose a more feasible way to assign new users to clusters, avoiding repetitive clustering.

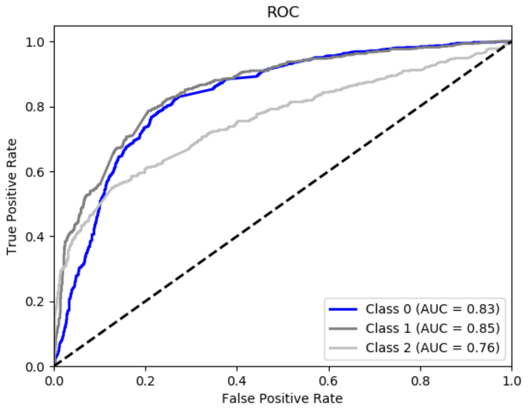


Figure 3.8: ROC curves, three categories.

Given that there is an existing categorization of user profiles, the feed-forward neural network presented in Section 3.3.2 is used to classify a new user into a suitable group. Once clusters have been obtained, we feed them as input to train the neural network and perform the testing using the ten-fold cross-validation procedure. It is worth mentioning that we use three clusters instead of four as explained in Section 3.2.2.

The final performance measured by means of ROC curves is depicted in Figure 3.8. In particular, the AUC values for Class 0, Class 1, and Class 2 are 0.83, 0.85, and 0.76, respectively. The curves representing the three classes reside near the upper left corner, implying a good prediction performance. Overall, the curves and the AUC values demonstrate that the obtained performance is much better compared to that before clustering in Figure 3.5 and Figure 3.6. This suggests that properly clustering user profiles can substantially increase the neural network's prediction performance.

Next, we validate the performance of PisaRec as follows. We opted for the ten-fold cross validation technique [103], where the dataset is split into ten equal folds, and the evaluation is done in ten rounds. By each round, one fold is utilized as testing, and the other nine folds are merged to create the training data. In a testing fold, for each user, the features are split into two parts, one part is fed as query, and the remaining part is removed to be used as ground-truth data. The ratio of the number of settings used as query to the total number of settings is called α. This simulates a real scenario, where the user already specified some settings, and the system is expected to recommend the rest, corresponding to the ground-truth data. For each user, PisaRec returns a ranked list of N settings (N is configurable), and the evaluation metrics are computed on the test set as follows. The recommended items are then compared with the ground-truth data to evaluate the performance. Eventually, we average out the metrics obtained from the testing folds to produce the final results.

We experiment with different configurations by varying α, k: the number of neighbor users used for the computation, N: the number of recommended items. In particular, $\alpha = \{0.1, 0.3, 0.5\}$; $k = \{3, 5, 10, 15\}$;

and N is varied from 1 to 50, simulating a real-world scenario where users have to set several settings. The precision-recall curves are then sketched following these parameters.

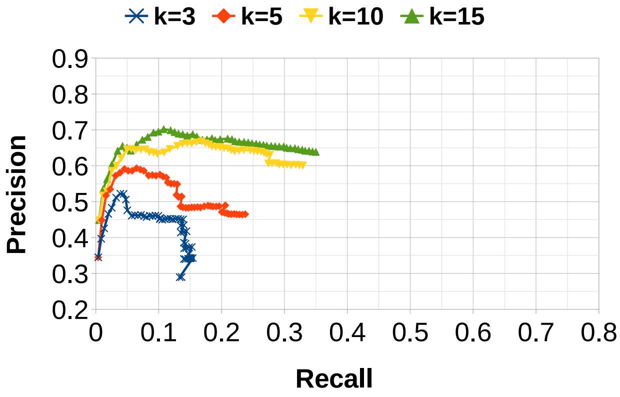


Figure 3.9: Configuration C_1.

As seen in Figure 3.9, when $\alpha = 0.1$, i.e., only a small amount of data is used as query, PisaRec recommends relevant settings to users, however with considerably low precision and recall. For instance, when $k = 3$, a maximum precision of 0.52 is obtained and the maximum precision is 0.7 when $k = 15$. Similarly, the recall scores are low, i.e., smaller than 0.4 by all the configurations. Altogether, this implies a mediocre performance which is understandable as the configuration with $\alpha = 0.1$ corresponds to the case where the user only specified a few settings, and the system has limited context to recommend additional settings.

When we increase α to 0.3, there is an improvement in both precision and recall as in Figure 3.10, compared to the results obtained with $\alpha = 0.1$ in Figure 3.9. Precision scores are always larger than 0.55 in the configurations, with 0.80 being the maximum value. Similarly, we also see that recall scores are gradually improved. For instance, a maximum recall of 0.35 is achieved with $k = 3$, and the corresponding maximum for $k = 15$ is 0.48.

Such an improvement is more evident when $\alpha = 0.50$, i.e., a half of the settings is used as query. In Figure 3.11, apart from some outliers, most of the precision scores are larger than 0.70, with 0.85 as the maximum value.

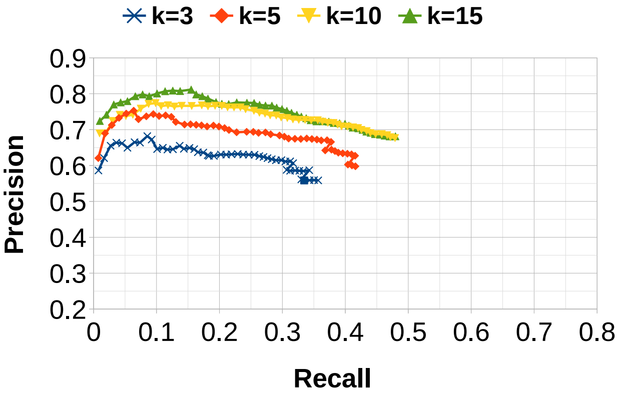


Figure 3.10: Configuration C_2.

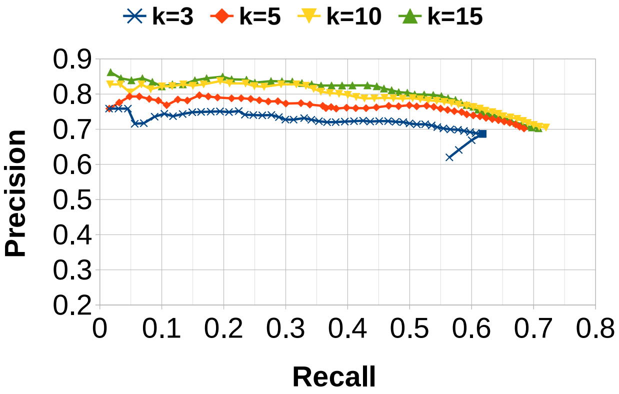


Figure 3.11: Configuration C_3.

Compared to the previous configurations with $\alpha = 0.1$ and $\alpha = 0.3$, the recall scores are also better, i.e., with a longer list of items, recall increases substantially. In particular, a recall of 0.73 is seen when $k = 10$ and $k = 15$.

Concerning the number of neighbors used for computing recommendations, i.e., k (see Section 3.3.3 and Formula 3.3), by considering Figure 3.9, Figure 3.10, and Figure 3.11 together, it is evident that adding more users for the computation contributes to a better prediction performance. For instance, by increasing k from 3 to 5, 10, and 15, we boost both precision and recall by all the cut-off values N.

Altogether, the experimental results show that even if users perceive their categories differently as shown in $\mathbf{RQ}_{2.1}$, once we have identified their right privacy group PisaRec can exploit the categories to provide relevant settings to users, though the considered dataset is pretty small. We anticipate that its performance can be further enhanced, if there is more data for training.

Answer to $\mathbf{RQ}_{2.3}$. PisaRec recommends highly relevant settings to a user, though there is limited amount of data available for training. The prediction performance improves alongside the amount of data fed as input.

3.5 Discussion

This section provides discussion related to the possible extensions of our work, as well as the threats to validity of our findings.

3.5.1 Extendability

Dataset. In our work, we utilized a relatively small dataset for the evaluation, the original data set contained 295 records for 437 features that we expanded to 444 through grouping. The amount of training data may impact the performance of both the clustering and classification phases. Moreover, as PisaRec is a collaborative-filtering recommender system, its performance is heavily driven by the quality and amount of data. We anticipate that we may need to calibrate the systems' parameters to maintain both timing efficiency and effectiveness with more data.

The unsupervised algorithm. In the scope of this chapter, we used the K-Medoids algorithm to cluster the user profiles. Such a technique has been chosen due to its simplicity and effectiveness. In fact, several clustering algorithms could be employed to categorize user profiles. Thus, the outcome of a clustering solution depends heavily on the considered techniques. We plan to extend our work by considering other clustering algorithms, such as CLARA [98], or DBSCAN [66].

The supervised classifier. The neural network used to classify user profiles may be suitable only for the considered dataset. For a different dataset, it is necessary to find adequate network configurations employing an empirical evaluation. For instance, the number of hidden layers, or the number of neurons for each layer, should be considerably increased to deal with a larger number of user profiles.

3.5.2 Threats to validity

We are aware of the existence of some threats that might harm the validity of the performed experiments as they are presented as follows.

- Threats to *construct validity* are related to any factor that can compromise the validity of the given observations. The main threat to construct validity is related to the size of the analyzed data. The used dataset is indeed relatively small but has the advantage of coming from a recent work [162] thus reflecting users' contemporary privacy behaviors. More extensive experiments are under planning encompassing other ethical dimensions beyond privacy.

- Concerning the threats to *internal validity*, i.e., any confounding factor that could influence our findings, we attempted to avoid any bias in the automatic creation of user profiles and in the way we split the full data into groups. We tried to mitigate this threat by semantically analyzing and double-checking the clusters obtained by the proposed approach.

- Concerning the threats to *external validity*, they are related to the generalizability of our results. This is about checking the adequacy of our privacy profiles in other contexts, notably, in the traveling or IoT domains. Generalizability is actually our initial driver for extracting privacy profiles from general moral questions. Thus, further experimental evidence is planned to support the reported chapter results.

3.6 Related work

The work presented in this chapter has been done in the context of the design of the EXOSOUL research project that aims at providing users with a personalized software layer that mediates users' interactions with the digital world according to user's ethics, including privacy preferences [158, 22].

According to various studies [128, 142], the vast majority of users do not bother to read privacy agreements because of the excessive language and confusing explanations [27, 96, 124, 157]; it is unreasonable also to expect they will read them on a regularly basis [123]. Resignation from privacy choices may also be a result of their dissatisfaction with the lack of options and excessive complexity [41].

Privacy profiling is at the core of our work therefore, most related studies are on user clustering, privacy profiling, and privacy preferences settings. The more significant part of existing studies about privacy profiling develop on the work of Westin [188]. Based on a series of privacy-related surveys, the author established "*Privacy Indexes*" for most of these polls to summarize results, indicate trends in privacy concerns, and suggest a widely recognized segmentation methodology of "*Privacy Profiles.*"The methodology he applied classifies people into three categories: *privacy fundamentalists, pragmatists, and unconcerned.* Because of the commercial nature of Westin's surveys, the methodology and the details of how privacy indexes were calculated are not fully disclosed, so we rely on subsequent works [106] that deeply analyzed and reported them.

Westin Segmentation, and particularly the pragmatism adherence of the consumers were criticized by the work of Hoofnagle and Urban [93, 183]. Their experimental work investigated customer expectations for privacy safeguards, showing that many people believe they have greater protection than they really do due to a lack of knowledge about corporate practices, privacy policies, and data usage limits. Westin's methodology has been applied, revised, and expanded in several empirical studies on privacy that include collected data. The categorizations that are most relevant to our research are reported in Table 3.1.

Dupree *et al.* [60] analyzed data from surveys and participants interviews, the authors identified five user clusters that emerge from end-user behaviors, including *Fundamentalists*, *Lazy Experts*, *Technicians*, *Amateurs* and the *Marginally Concerned.*

Schairer *et al.* [165] presented a model of privacy disposition and its development based on qualitative research on privacy considerations in the context of emerging health technologies. The authors identified six clusters, including *Fatalism*, *Nothing to hide*, *Something to hide*, *Tradeoff*, *Personal responsibility*, *Moral right.*

In the research proposed by Sanchez *et al.* [162] the authors presented the results of a fitness-related simulation and questionnaire to classify users according to their privacy-related preferences. They used two different sets of labels for their clusters, one for the computed privacy-profile assignment consisting of six groups and one for self-assessment they proposed to the users consisting in four groups. The first clusters were labeled as: *Unconcerned*, *Socially Active*, *Health-focused*, *Minimal*, *Anonymous*, *Strict.* The second clusters were labeled as: *Privacy Conservative*, *Unconcerned*, *Fence-Sitter*, *Advanced User.*

Chapter 4

Developing a user's ethical profile: survey design and administration

New ethical concerns, beyond those related to data protection and privacy violations, have arisen as a result of the development and proliferation of digital technologies that are becoming increasingly autonomous in our society. Users are not protected in their interactions with digital technology, but at the same time, autonomous systems are free to dominate the realm of decision-making that is traditionally reserved for human beings. Within this framework, the Exosoul project is an interdisciplinary endeavor with the goal of constructing a personalized software exoskeleton that mediates behaviors in the digital world according to the moral choices of the user. The exoskeleton relies on ethical profiling of users, which has a purpose comparable to that of privacy profiling presented in the literature but aims to reflect and forecast general moral preferences instead. Our strategy is a hybrid one that begins with the identification of profiles in a top-down fashion and then moves on to the refinement of profiles through an approach that is personalized and data-driven. In this chapter, we discuss the results of our preliminary experiment on the construction of top-down profiles. We take into consideration the correlations between different ethical positions (such as idealism and relativism), personality traits (such as honesty/humility, conscientiousness, Machiavellianism, and narcissism), and worldviews (such as normativism), and then we use a clustering approach to create ethical profiles that are predictive of user's digital behaviors regarding privacy violation, copy-right infringements, caution, and protection.

4.1 Background

There is a general and growing consensus that the diffusion of autonomous digital technology may harm the values our societies are based on. Europe is at the forefront of the elaboration on these issues especially from a regulatory point of view. First concerning privacy with the GDPR [149] and currently with the AI act [42]. Regulation is important and represents the level of awareness that the society as a whole has matured regarding the potential misuse of digital technologies. However it is not sufficient and cannot cover all the space of potential misuses concerning the risk which attains the core of the fundamental rights of the citizens. Privacy concerns are insufficient: ethics and the human dignity are at stake [94, 95]. As a matter of fact, individuals are unprotected and powerless in their interaction with the digital world. In a digital society where the relationship between citizens and machines is uneven, moral values like individuality and responsibility are at risk. Despite the ideal of a human-centric AI and the recommendations to empower the users, the power and the burden to preserve the users' rights still remain in the hands of the (autonomous-) systems producers.

Recent studies have claimed the importance of automating the decisions concerning privacy and digital behavior by using forms of recommendations including personal assistant agents. For example, Kola *et al.*. [105] aim at building a personal assistant agent on the basis of different factors, including the Schwartz's moral values [168], in order to model the social situation of a user and act accordingly to her "priority".

In the privacy domain, privacy personal assistants were proposed in Sadeh *et al.*. [161] and in Ulusoy *et al.*. [182]. More in general several approaches tried to help the user with respect to privacy and security when interacting with digital means [10].

The EXOSOUL [22] project aims to empower humans with an automatically generated exoskeleton, i.e., a software shield that protects them and their personal data through the mediation of all interactions with the digital world that would result in unacceptable or morally wrong behaviors according to their ethical and privacy preferences. The exoskeleton relies on the ethical profiling of a user, similar in purpose to the privacy profiling proposed in the literature [165, 61], but aiming at reflecting general moral preferences and predicting user's digital behaviors accordingly like proposed in [127]. More precisely, EXOSOUL is based on the notion of digital ethics [71] and its separation in soft ethics to reflect user's ethics and hard ethics to define the ethical values a digital system shall comply with. Our aim is to develop ethical profiles, by which to predict digital behaviors. Thus, in line with the EXOSOUL hybrid approach, the first step is the construction of such profiles in a top-down manner. For this reason, here come our research questions:

RQ3.1: *Is it feasible to create a user profile that takes into account their ethical stances?*

RQ3.2: *Are the ethical profiles helpful in predicting how people would behave online?*

4.2 Contextualization of the Theory

In order to define users" ethical profiles we considered the Ethics Position Theory (EPT) [74, 77, 147], which suggests individuals" differences in moral judgments. This theory posits that the individuals" personal morality contains unique and idiosyncratic elements that depend on the experience related to moral issues. These elements are sustained by two dimensions: idealism and relativism. The former reflects absolute moral principles, mostly oriented to truth, benevolence and avoiding harming others. The latter relies on the careful evaluations of the situations, contexts and consequences, and also reflects the idea that harm is sometimes necessary in order to produce good. These two dimensions are based on the deontological and teleological models, respectively. The deontological perspective defines the rightness and wrongness of possible courses of action, by comparing them with pre-determined norms representing personal values, whereas the teleological perspective considers the perceived consequences, the probability and the desirability of each course of action for various stakeholder groups, and the importance of each group. The EPT does not assume that individuals are only rule-oriented (idealism) or consequence-oriented (relativism), but rather assumes that individuals can range from high to low in their idealism and relativism ideologies. This leads to identify four moral philosophies: situationism, subjectivism, absolutism, and exceptionism (see Table 4.2, [147]).

	Low Relativism	**High Relativism**
Low Idealism	**Exceptionists**: Conventionalists who tolerate exceptions to moral standards when benefits offset potential harmful consequences.	**Subjectivists**: Realists who do not endorse moral standards that define right and wrong or the avoidance of harmful consequences
High Idealism	**Absolutists**: Principled idealists who endorse both reliance on moral standards and striving to minimize harm done to others	**Situationists**: Idealistic contextualists who value minimizing harm rather than reliance on moral standards that define right and wrong

Table 1. The four moral types identified in EPT.

The EPT has been related to personality, which includes dispositions, temperaments, characters, attitudes, values, and so forth. Personality is "the dynamic organization within the individual of those psycho-physical systems that determine his unique adjustments to the environment" (p. 48) [18], by creating "....the person's characteristic patterns of behavior, thoughts, and feelings" (p. 48) [17]. Although the personality attributes are independent of each other, some can align to form profiles of shared similarities. This means that individuals' ethics positions are systematically related, amongst others, to personality [75]. The most consolidated approach to personality assumes that a small number of enduring dispositions or traits provides the basis for differences and similarities among people. In this vein, the Five Factor Model (FFM) [51, 52] is the most useful framework, given that the 5 personality traits identified have emerged with great regularity in different studies: neuroticism, extraversion, conscientiousness, agreeableness, and openness. Afterward, the Hexaco theory of personality identified a sixth factor: honesty/humility trait [109, 110]. In addition, the construct of Dark Triad personality was also proposed, based on the idea that the human nature encompasses also a constellation of sub-clinical and malevolent traits [150], which reflect different facets of antisocial and adverse personality [153]: psychopathy, Machiavellianism, and narcissism.

In particular, honesty/humility and conscientiousness traits were found to be mostly associated positively with morality and integrity [38, 152], and highly resistant to moral disengagement [39, 145, 57, 130]. Basically, these two personality traits can give insights into what a moral character is [37].

Specifically, the honesty/humility trait reflects sincerity, fairness, greed, modesty, and no manipulation, no interested in richness, luxuries, and elevated social status. This trait is characterized by "the tendency to be fair and genuine in dealing with others, in the sense of cooperation with others even when one might exploit others without suffering retaliation" (p. 156) [20]. It was found related positively to individualizing values (e.g., sensitivity to harm and fairness) [198]. This leads people with high honesty/humility to conform to idealism rather than to relativism. Indeed, previous studies showed that honesty-humility is correlated positively to idealism and negatively to relativism [39, 145].

Conscientiousness reflects the tendency of being persistent in the pursuit of goals/tasks, and relies on orderliness, meeting of obligations, self-confidence, self-regulation and self-discipline, integrity, fairness, and inhibition of impulses that go against moral obligations. Conscientious people can determine their ethical codes, are less social and tend to sacrifice social ties for achieving goals [122]. Conscientiousness was related to the exceptionist moral philosophy, being not correlated either to idealism or relativism [99]. However, other studies showed that this trait was either positively correlated to relativism [75], or positively to idealism but negatively to relativism [39, 145].

As concerns the Dark Triad construct, the three malevolent personality traits were found related to moral disengagement and *schadenfreunde* (feeling of pleasure at another's suffering) [65]. Notably, Machiavellianism and psychopathy are highly correlated, especially in relation to the Big Five personality traits. Thus, these two personality traits are seen as a single psychopathic entity [143], whereas narcissism is considered the lightest dimension of the Dark Triad [80]. Besides, these two traits are more predictive of digital behaviors. Machiavellianism and narcissism were found positively related to self-promotion on Facebook [160], to the number of personal information disclosed online [163, 155], and to the tendency to self-disclose in computer-based communication [163, 155]. Machiavellianism and vulnerable narcissism predicted also problematic social media use [100] and less congruence between the true self and the Instagram self [83].

In details, Machiavellianism reflects actions oriented to pragmatism, entails a lack of affect in interpersonal relationships, and moral standards. High Machiavellianism is associated with manipulative and cynical tactics, the basic attitude being "the ends justify the means' [34, 40]. Machiavellianism is characterized by an utilitarian rather than a moral view when interacting with others [34]. Different studies showed that Machiavellianism correlates negatively to idealism but positively to relativism [39, 107, 125]. Following Leary *et al.*. [107], Machiavellianism is basically associated with the subjectivism moral philosophy.

Narcissism is supported by an inflated opinion of self, feelings of entitlement, superiority, need for admiration [64], egoism and arrogance [76]. It is also characterized by moral disengagement [145], unethical attitudes and questionable behaviors aimed at pointing out high achievements, which in turn reiterate admiration, power, and ego superiority [84, 151]. Narcissism was found to correlate positively to idealism but not to relativism [184], or to be higher in relativistic egoists cluster than idealistic altruists cluster [35]. Yet, the narcissistic gratification dimensions correlated to idealism but not to relativism [174]. These contradictory results support the idea that narcissists can behave morally "right" to increase their ego and get higher levels of narcissistic supplies from helping others, even though they are morally disengaged.

Notably, beside personality, individuals' ethics positions can be also related to worldviews. Following the Polarity Theory [179], human worldviews are defined by humanism, which relies on humanity and experiences as intrinsically valuable, and normativism, which relies on values determined by external norms and ideals [134]. Thus, humanistic people rarely engage in misconduct because they hold group values and beliefs and seek affiliated interests [172]. Humanism was found related positively to the political left, preferences for equality, openness, emotionality and honesty/humility traits, as well as to moral intuitions pertaining to fairness and prevention of harm, and negatively to levels of authoritarianism, social dominance, general and economic system justification [137]. By contrast, normativism is related positively to the political right, conservative issue preferences, resistance to change, acceptance of inequality, and negatively to openness, emotionality and honesty/humility traits [135], as well as with moral intuitions pertaining to ingroup loyalty, respect for authority, and protection of sanctity [137]. Given this picture, it is reasonable to assume that humanism is more compatible with idealism, whereas normativism with relativism. However, normativism could be also associated with idealism if people idealize their worldview. In the present study we used only the normativism worldview scale [136] because the humanism worldview resembles the honesty/humility personality trait.

Based on the literature reviewed above, the ethics positions (idealism and relativism), personality (honesty/humility, conscientiousness, Machiavellianism and narcissism) and worldview (normativism) variables were combined by a clustering approach in order to identify ethical profiles.

4.3 Methods

4.3.1 Participants

330 participants (182 females, 138 males and 10 individuals with no gender information; mean age = 20.52 years, s.d. = 2.55; age range = 18-35 years) were recruited from the University of L'Aquila, Italy. Participants were not requested to give their informed consent, given that the

questionnaire was anonymous, in compliance with the European privacy legislation GDPR. The internal review board approved the study.

4.3.2 Materials

The following tests were administered.
The Italian adaptation of the Ethics Position Questionnaire (EPQ-5) [147]: 5 items for idealism (e.g., A person should make certain that their actions never intentionally harm another even to a small degree "also actions performed by Internet or the computer') and 5 for relativism (e.g., What is ethical varies from one situation and society to another "also situations that occur on the Web'). Given the purpose of the present research, beside each item, we reported an example of an action referred to the digital world (the examples are not included in the original questionnaire). For each item, participants were asked to rate the degree of agreement from 1 (strongly disagree) to 5 (strongly agree). The item analysis showed that the corrected item-total correlation was acceptable for both scales (greater or equal to .30 [141]). The reliability (internal consistencies) was $\alpha = .63$ for the idealism scale, and $\alpha = .69$. for the relativism scale.
The Italian adaptation of the HEXACO-60 [21]: 10 items for honesty/humility (e.g., I wouldn't use flattery to get a raise or promotion at work, even if I thought it would succeed) and 10 for conscientiousness (e.g, I plan ahead and organize things, to avoid scrambling at the last minute). For each item, participants are asked to rate the degree of agreement from 1 (strongly disagree) to 5 (strongly agree). The item analysis showed that 3 items for each scale were not satisfactory in terms of corrected item-total correlation [141]. Thus, 7 items were retained for each scale. The reliability was $\alpha = .68$ for the honesty/humility scale, and $\alpha = .69$ for the conscientiousness scale.
The Italian version of the Dark Triad Dirty Dozen [167]: 4 items for Machiavellianism (e.g., I tend to manipulate others to get my way) and 4 for narcissism (e.g., I tend to want others to admire me). For each item, participants were asked to rate the degree of agreement from 1 (strongly disagree) to 5 (strongly agree). The item analysis showed that the corrected item-total correlation was acceptable for both scales. The

reliability was $\alpha = .77$ for the Machiavellianism scale, and $\alpha = .81$ for the Narcissism scale.

The Italian adaptation of the normativism scale [136]: only 3 items were considered from the "political values" sub-section (e.g., The maintenance of law and order is the most important duty of any government). For each item, participants are asked to rate the degree of agreement from 1 (strongly disagree) to 5 (strongly agree). The item analysis showed that all items were satisfactory in terms of corrected item-total correlation. The reliability was $\alpha = .72$.

The digital behavior scale: 3 items for privacy violation (e.g., I use the personal information of others without permission - e.g., a photo of a friend) and 3 items for copy-right infringements (e.g., I use software without owning a licence) were adapted in Italian from the Unethical Computer Using Behavior Scale (UCUBS) [132]; Then, 4 items adapted in Italian from Buchanan *et al.*. [29] were used to measure caution (e.g., Do you read a website's privacy policy before registering your information?). For each item, participants were asked to rate the degree of frequency of each action from 1 (never) to 5 (always). The reliability was $\alpha = .66$ for privacy violation, $\alpha = .66$ for copy-right infringements, and $\alpha = .70$ for caution.

In general, all reliability coefficients are satisfactory (see [175]).

4.3.3 Procedure

The questionnaire was administered by means of the platform "LimeSurvey". The address of the questionnaire was disseminated to the students in the the form of a URL and by a QR-Code to facilitate its use by camera-equipped devices. To avoid the order effect due to fixed order of the questions, the scales were presented randomly across participants. The questionnaire administration lasted about 15 minutes.

4.3.4 Plan of Analysis

Statistical analyses were performed by IBM SPSS Statistics. Data were transformed in z-scores. The standard cut-off of ±3 standard deviations away from the mean was used to remove outliers: in total 13 outliers were detected and excluded from subsequent analyses. This led to a final sample of 317 participants (177 females, 130 males, and 10 individual with no gender information; mean age = 20.53 years; s.d. = 2.58). First of all, the cluster analyses were carried out using as variables idealism, relativism, honesty/humility, conscientiousness, Machiavellianism, narcissism and normativism. Two different approaches were pursued. The first approach was aimed at exploring the dataset using the two-step method, that is conducting a hierarchical cluster analysis using Ward's method, and a subsequent confirmatory k-means analysis. This method provides a relatively robust identification of clusters [178]. The second approach was and aimed at confirming the ETP, which identifies four moral philosophies: situationism, subjectivism, absolutism, and exceptionism. Thus, only the k-means confirmatory analysis was used, with 4 predetermined clusters. For each approach, univariate analyses of variance (Anova) were conducted to explore the validity, using idealism, relativism, honesty/humility, conscientiousness, Machiavellianism, narcissism and normativism as dependent variables, and the cluster solution as the independent variable. In addition, the discriminant analysis was performed to clarify the classification of the cases according to the cluster solution. Finally, Anovas were also conducted to explore the predictive power of the cluster solutions in terms of digital behavior.

4.4 Results

4.4.1 Two-step cluster analysis - Hierarchical and k-means

The cluster solution was determined considering the dendrogram, the agglomeration schedule coefficients, and the interpretability of the clusters solution [14]. The analysis suggested a 2-cluster solution (see Figure 4.1).

The k-means analysis was carried with $k = 2$ (see Figure 4.2). The silhouette value, which is a measure of how similar an object is to its own cluster compared to the other clusters, is 0.3, which is considered

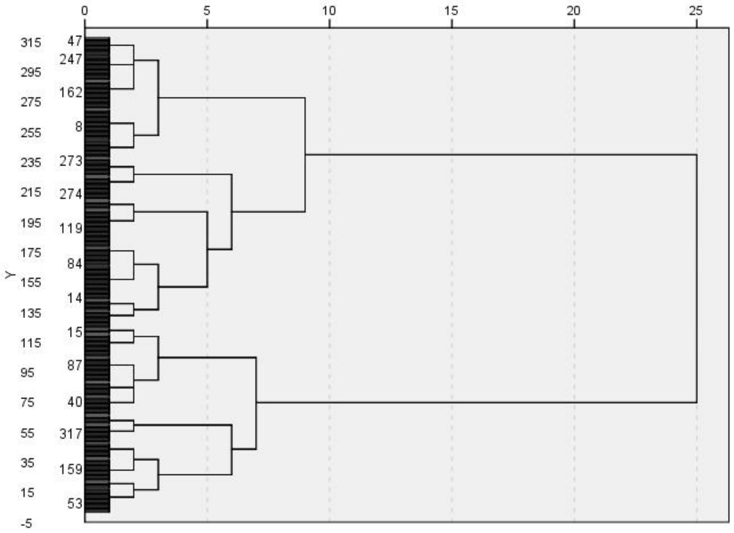


Figure 4.1: Dendrogram using the Ward's method

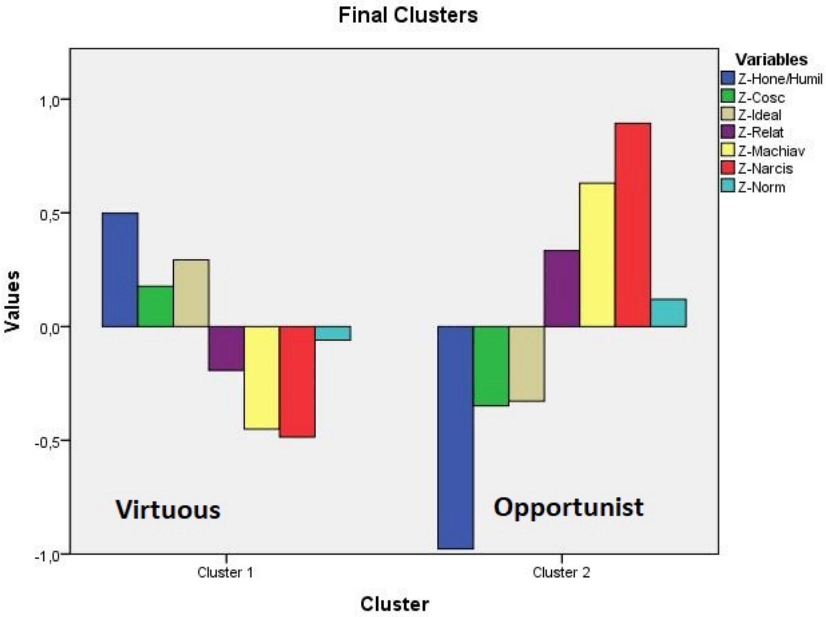


Figure 4.2: 2-Cluster solution

acceptable. The first cluster was formed by 210 subjects, whereas the second cluster by 107 subjects. Based on the variables distribution within each cluster, we defined the first cluster "virtuous" and the second cluster "opportunist".

Regarding the differences between the input variables of the cluster solution, the "virtuous" cluster produced higher scores in idealism, honesty/humility, conscientiousness, and lower scores in relativism, Machiavellianism and narcissism than the "opportunist" cluster. No difference was found between the two clusters in normativism (see Table 4.2).

Regarding the discriminant analysis, 98,7% of original grouped cases and 98.1% of cross-validated grouped cases were correctly classified.

	Virtuous: Mean (SD)	**Opportunist: Mean (SD)**	**Statistics: Anova**
Idealism	.293 (.055)	-.328 (.077)	F (1,315) = 43.25, p < .0001
Relativism	-.193(.066)	.333(.093)	F (1,315) = 21.16, p < .0001
Honesty/Humility	.498(.049)	-.978(.069)	F (1,315) = 300.85, p < .0001
Conscientiousness	.177(.067)	-.348(.094)	F (1,315) =20.82, p < .0001
Machiavellianism	-.450(.048)	.631(.067)	F (1,315) = 174.73, p < .0001
Narcissism	-.486(.050)	.894(.071)	F (1,315) = 253.66 p < .0001
Normativism	-.060(.069)	.120(.097)	F (1,315) = 2.290, p = .13

Table 4.2: Differences between the 'virtuous' and 'opportunist' clusters

As concerns the differences between the two clusters in the digital behaviour, the "virtuous" cluster scored lower in privacy violations [F (1,315) = 26.09, p < .0001; Mean(SD) -.197(.066) vs .387(.093)] and copyright infringements [F (1,315) = 29.89, p < .0001; Mean(SD) -.210(.066) vs .412(.093)], and higher in caution [F (1,315) = 6.37, p < .05; Mean(SD) .100(.068) vs -.197(.096)] than the "opportunist" cluster.

4.4.2 The cluster analysis - k-means

The cluster solution based on the four moral philosophies of the EPT also showed an acceptable silhouette value (.20) (see Figure 4.3).

The first cluster was formed by 86 subjects, the second by 79 subjects, the third by 77, and the fourth cluster by 75. Based on the variables distribution within each cluster, we defined the first cluster "legalist'; the second cluster "sensible'; the third cluster "opportunist'; the fourth cluster "virtuous". Regarding the differences between the input variables of the cluster solution, the "legalist", "sensible" and "virtuous" clusters showed no differences in terms of Machiavellianism and narcissism; the "legalist" and "virtuous" cluster did not differ in normativism; the "legalist cluster did not differ in idealism with respect to the "sensible" and "virtuous" clusters and in relativism with respect to the "opportunist" cluster; then, the "sensible" and "virtuous" cluster showed no difference

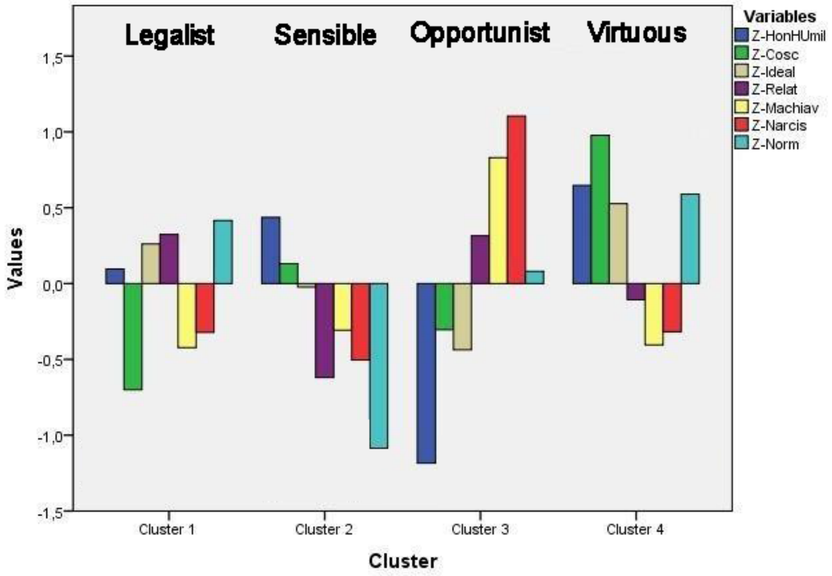


Figure 4.3: 4-Cluster solution

in terms of honesty/humility; the other comparisons were significant (see Table 4.3).

Regarding the discriminant analysis, %96.8 of original grouped cases and %95.3 of cross-validated grouped cases were correctly classified.

	Legalist Mean (SD)	**Sensible Mean (SD)**	**Virtuous Mean (SD)**	**Opportunist Mean (SD)**	**Statistics Anova**
Idealism	.261 (.083)^{a}	-.023 (.087)^{b}	.527 (.089)^{bc}	-.437 (.088)^{abc}	F (3,313) = 21.81, p < .0001
Relativism	.324 (.099)^{a}	-.619 (.103)^{ab}	-.106 (.106)^{abc}	.315 (.105)^{bc}	F (3,313) = 18.87, p < .0001
Honesty/Humility	.095 (.077)^{a}	.437 (.081)^{ab}	.647 (.083)^{ac}	-.1184 (.082)^{abc}	F (3,313) = 100.75, p < .0001
Conscientiousness	-.700 (.085)^{a}	.131 (.088)^{ab}	.978 (.091)^{abc}	-.304 (.090)^{abc}	F (3,313) = 65.91, p < .0001
Machiavellianism	-.423 (.074)^{a}	-.308 (.077)^{b}	-.405 (.079)^{c}	.830 (.078)^{abc}	F (3,313) = 61.28, p < .0001
Narcissism	-.322 (.080)^{a}	-.504 (.083)^{b}	-.318 (.086)^{c}	1.105 (.085)^{abc}	F (3,313) = 79.05, p < .0001
Normativism	.415 (.082)^{a}	-1.086 (.086)^{ab}	.589 (.088)^{bc}	.080 (.087)^{abc}	F (3,313) = 76.78, p < .0001

[*] Significant differences between clusters are showed by letter correspondences; Tukey's Post-hoc (HSD) were used.

Table 4.3: Differences between the "legalist", "sensible", "virtuous" and "opportunist" clusters

As concerns the differences between the four clusters in the digital behaviour, the "legalist", "virtuous" and "sensible" clusters showed no difference in privacy violation; the "opportunist" cluster showed higher scores than the other 3 clusters. Regarding copy-right infringements, the "legalist" cluster showed no difference as compared to the "sensible" and "virtuous" clusters; the "sensible" cluster showed higher scores than the "virtuous" cluster; then, the "opportunist" cluster showed higher scores than the other 3 clusters. Regarding caution, the "virtuous" cluster showed higher scores only than the "opportunist" cluster; the other comparisons were not significant (see Table 4.4).

	Legalist Mean (SD)	**Sensible Mean (SD)**	**Virtuous Mean (SD)**	**Opportunist Mean (SD)**	**Statistics Anova**
Privacy violation	-.065 (.104)^{a}	-.150 (.109)^{b}	-.246 (.112)^{c}	.466 (.110)^{abc}	F (3,313) = 8.35, p < .0001
Copy-right infringements	-.097 (.104)^{a}	.034 (.108)^{b}	-.374 (.111)^{bc}	.438 (.110)^{abc}	F (3,313) = 9.41, p < .0001
Caution	.029 (.106)	.048 (.111)	.226 (.114)^{a}	-.302 (.112)^{a}	[F (3,313) = 3.80, p < .05

* Significant differences between clusters are showed by letter correspondences; Tukey's Post-hoc (HSD) were used.

Table 4.4: Differences between the "legalist", "sensible", "virtuous" and "opportunist" clusters

4.5 Discussion of the profiling results

Creating ethical profiles that are capable of being predictive of digital activities was the goal of this section. Given that they were formed in part on the basis of personality qualities, which normally do not change over the course of years, these ethical profiles may be deemed to have a good chance of being relatively stable over time. To be more specific, ethical profiles were developed using two distinct clustering solutions (k = 2 and k = 4), which combined the Ethics Position Theory (idealism and relativism) with personality (honesty/humility, conscientiousness, Machiavellianism, and narcissism) and worldview (normativism) variables. Normativism was used as the criterion for determining whether or not an individual adhered to societal norms.

The investigations demonstrated that the 2-cluster solution, which is based on the interpretation of the dendrogram, is an actionable one. This is due to the fact that it demonstrates an overlap between the clusters exclusively in terms of normativism. In addition, the discriminant analysis demonstrated that the instances that were initially grouped together were properly categorized a significant proportion of the time. The "virtuous" cluster has a greater propensity to commit fewer violations of copyright and privacy laws in comparison to the "opportunist" cluster. This lends credence to the idea that the "virtuous" cluster is governed by *principled principles* (a set of ethical values and moral standards that guide an individual's or group's behavior), in contrast to their opponent, which is more likely to break norms in order to further their own personal interests. This conclusion expands prior results about the ethical profiles in marketing discussions to include digital actions [13]. In addition, the "virtuous" cluster received a better score than the "opportunist" cluster in the categories of privacy setting and information registration on the internet (caution). This would imply that the "virtuous" cluster is more attentive and careful than the "opportunist" cluster since they pay greater attention to detail.

The concept that the four moral philosophies of the EPT might drive the combination of a person's personality characteristics and their normativistic world view is the foundation for the 4-cluster solution. The findings indicated that there was considerable overlap between the "legalist", "virtuous", and'sensible" clusters, particularly in terms of Machiavellianism and narcissism. This was the case despite the fact that the cases that were originally grouped were accurately categorized with a substantial majority. It is likely that these two personality qualities grouped together due to the fact that they are components of the Dark Triad construct. As a result, there is overlap between the three personality traits clusters. The "opportunist" cluster was the only one that exhibited greater scores than the other three groups in terms of privacy violations and copy-right infractions. Only with regards to copyright infringements did the "legalist" cluster exhibit lower ratings than the "opportunist" cluster, which suggests that legalists pay attention to this topic just as much as reasonable and virtuous individuals do. However, the "virtuous" cluster was the only one to have lower scores in caution than the "opportunist" cluster.

When considered as a whole, these findings imply that the two alternative cluster solutions produce distinct outcomes not just as a consequence of the inherent ethics-, personality-, and worldview-related characteristics, but also as a function of the digital behavior that was investigated (privacy violation, copyright infringement and caution). Importantly, these results emphasize that ethical profiles may be constructed as a mix of numerous components, including personality characteristics that contain ethical features. This is an important discovery since it demonstrates that ethical profiles can be built. In this regard, it is noteworthy to highlight that ethical profiles may be characterized in both good and negative terms, each of which has evident repercussions for one's digital activity. The latter in this research indicated largely hostile attitudes and caution, although of course it would be vital to analyze digital behaviors more accurately in terms of general privacy (for example, disclosure of information) and security knowledge.

Given that it depicts two ethics that are diametrically opposed to one another, the 2-cluster solution may, in general, serve as a basis for the development of more refined ethical profiles (virtuous vs opportunist). Even if it only partly corroborated the EPT, the 4-cluster solution is still somewhat of a work in progress, and it has to be tested by separating the clusters more clearly.

4.6 Conclusions on profiling results

These findings are very promising and open many directions for refinements. Future works are aimed at better disentangling not only the "legalist", "sensible" and "virtuous" clusters, but also the "opportunist" cluster in sub-clusters. This might be achieved by considering other variables. For example, one can include other scales measuring integrity, autenticity, dogmatism, moral disengagement, sadism, opportunism, normlessness, and so forth. In terms of personality variables one can add neuroticism and openness to experience. Referring to the "virtuous" and "opportunist" clusters, it would be interesting to clarify if they can be unpacked in more specific clusters. Indeed, we used mainly moral variables and morally oriented personality traits. Variables that are more oriented to measure normative and consequential ethics, such as legalism, rationality, utilitarianism, authority, and so forth, could also be explored. In addition, we intend to study the same research questions using machine learning algorithms. We are interested in investigating how machine learning algorithms perform in predicting the digital behaviour of users from ethics positions and personality characteristics.

Finally, the findings provided by this study could be also generalized to other digital behaviors, such as those related to social networks. Notably, the sample used in the present study was formed by university students, which limits the applicability of the results to the larger population present in the social media. Therefore, future studies should also explore the key role of ethical profiles in predicting digital behaviors in different age samples (e.g., aged people), with different educational levels and cultural backgrounds.

Chapter 5

Application scenario: profile driven cookies management

5.1 Background

In this chapter, we are going to analyze a typical scenario where a user is proposed to express choices about the impact of cookie management while visiting a website. We will perform the analysis in light of the GDPR directive and the characters, responsibilities, processes, stakeholders and rules set forth therein. Other than a brief overview on web privacy problems, we will also focus on how EXOSOUL could automatically manage cookies taking advance of a user's profile.

5.2 Legislation, the European example: GDPR

The General Data Protection Policy (GDPR) is a regulation that addresses all aspects of data privacy and data security. It is widely acknowledged as having the strictest requirements anywhere in the world. Regardless of the fact that the European Union (EU) was responsible for developing and enforcing the rule, it is applicable to businesses based anywhere in the globe if such businesses target persons residing in the EU or collect data connected to citizens of the EU. On May 25, 2018, the legislation went into effect, and since then, businesses that breach its privacy and security standards risk significant penalties. These penalties include fines that might reach tens of millions of euros if they are not paid. The General Data Protection Regulation (GDPR) sends a powerful signal that the European Union (EU) is adopting a demanding position on data privacy and security in an age in which people are increasingly entrusting their personal data to cloud services and frequent data breaches occur. Compliance with the General Data Protection Rule (GDPR) may be a daunting undertaking, especially for small and medium-sized organizations, due to the complexity, breadth, and sometimes need for more definition within the regulation. [195]

5.2.1 History of the GDPR

The European Convention on Human Rights, established in 1950, contains a provision that recognizes the right to privacy. Specifically, this clause states that "Everyone has the right to respect for his private and family life, his home and his correspondence." This provision is considered to be an integral component of the right to privacy. In response to the rapid advancements in technology and the proliferation of the Internet, the European Union (EU) has taken steps to ensure the protection of this fundamental right. The EU first recognized the need for data privacy and security measures in 1995 with the adoption of the European Data Protection Directive. This directive established basic data privacy and security requirements, which were then incorporated into the national legislation of each EU member state. However, as the Internet continued to evolve, it became increasingly clear that stronger protections were needed. With the advent of banner advertisements on the World Wide Web in 1994 and the widespread availability of internet banking by 2000, the EU began to consider the need for a more comprehensive approach to personal data protection. In 2011, a lawsuit was filed against Google alleging that the company had scanned emails, and two months later, Europe's data protection body called for the development of a comprehensive strategy for personal data protection. As a result, efforts were initiated to amend the 1995 directive. After being adopted by the European Parliament in 2016, the General Data Protection Regulation (GDPR) came into effect and became mandatory for all companies as of May 25, 2018. The GDPR represents a significant step forward in ensuring the protection of personal data and privacy rights, especially in light of the vast amounts of personal information being collected by companies in today's digital landscape. [195]

5.2.2 Mission, sanctions, and definitions

The General Data Protection Regulation (GDPR) regards any organization, regardless of its geographical location, that processes the personal data of all the individuals residing in the European Union (EU) or offers products or services to citizens of the EU. The General Data Protection Regulation (GDPR) establishes stringent duties on enterprises to safeguard the confidentiality and safety of personal data. If an organization is found to be in violation of the requirements outlined in the GDPR, they will be liable to harsh penalties, including fines of up to €20 million or 4% of their total worldwide revenue, whichever is greater. The General Data Protection Regulation (GDPR) gives to the individuals the right to sue for financial compensation in the event that they incur losses as a consequence of a data breach. In order for enterprises to be in full compliance with the GDPR, they need to first have an understanding of the GDPR's penalties and then take the steps required to protect themselves from being subject to those fines.

A wide variety of legal terminology are discussed at detail within the GDPR:

- **Personal data** — When discussing privacy and data protection, the term "personal data" refers to any information that, whether directly or indirectly, may be used to identify a specific person. This category of information can be gathered in a variety of formats, such as information that is directly identifiable, such as names and email addresses, as well as information that is more nuanced, such as location data, ethnicity data, gender data, biometric data, religious beliefs, online cookie data, and political opinions. These can all be collected. It is important to consider that even data with a pseudonym might be regarded as personal information if it is not difficult to identify the specific individual to whom the data belongs.

- **Data processing** — A data processing operation may be defined as any action made on data, regardless of whether the activity was performed manually or via an automated process. This encompasses a wide variety of actions, such as gathering, recording, organizing, constructing, storing, using, and erasing data. In its most basic sense, this involves almost every facet of the processing and administration of data.

- **Data subject** — The specific person whose information is being processed at this time. These are either customers you already have or those who have visited your website.

- **Data controller** — The person who is accountable for deciding the purposes of and methods for the processing of personal data. This category includes those who own a business or who work for that business and manage data in any capacity.

- **Data processor** — Whoever handles personal data on behalf of another data controller is considered a data processor. The General Data Protection Regulation (GDPR) imposes extra duties on the following individuals and organizations. [195]

5.2.3 Privacy rights for data subjects

The General Data Protection Regulation (GDPR) recognizes a wide variety of new privacy rights for people in relation to their own personal data. Individuals will have more control over the information they supply to organizations as a result of these rights, which are designed to provide people with more agency. In order to guarantee that they are in compliance with the GDPR, it is indispensable for enterprises to have a thorough grasp of these rights.

Here's an overview of the data subjects' privacy rights:

- **The right to be informed**

- **The right of access**
- **The right to rectification**
- **The right to erasure**
- **The right to restrict processing**
- **The right to data portability**
- **The right to object**
- **Rights in relation to automated decision making and profiling** [195]

5.3 Cookies and The Transparency & Consent Framework

5.3.1 What is a HTTP cookie

Web cookies, also known as Internet cookies, browser cookies, or just cookies, are small data files that are created by a web server and kept on a user's computer or mobile device. Cookies go by a variety of different names, including Internet cookies, browser cookies, or simply cookies. The intention of these cookies is to improve the user's browsing experience by enabling the website to retain information about the user's browsing activities. This information may include the user's previous visits to the site, items added to a shopping cart, and data submitted through forms. This information may include personally identifying information such as the user's name, address, and credit card information.

Cookies are an essential part of the user experience online because of the critical part they play in the upkeep of the user's stateful information and because of the fact that browsing the web requires them. They allow web servers to store information about the user's activities on the website and allow the site to retain specific information, such as the contents of a shopping cart, as well as monitor the user's browsing behavior, such as logging in, clicking particular buttons, and tracking previous visits to the site. Additionally, they allow the site to retain specific information, such as the contents of a shopping cart. In addition, cookies may be used to store sensitive information that the user has supplied through forms. This information might include email addresses, passwords, and credit card numbers.

Authentication cookies are often used by web servers as a way of confirming the identity of a user and the account that is being used for the purpose of logging in. Users would be needed to check in manually each time they accessed a website containing private information if these cookies weren't being used. The encryption of authentication cookies' contents, the safety of the website that supplies them, and the security settings of the user's web browser all contribute to the cookies' overall level of safety. However, if an attacker manages to get access to the contents of a cookie, this might result in the disclosure of sensitive information or even the user's credentials for the website that issued the cookie. This is because the cookie contains a hash of the user's information.

It is usual practice to employ *tracking cookies* in addition to authentication cookies placed by third parties in order to keep track of a user's online activity over the course of time. Concerns have been made over an invasion of privacy as a result of the usage of these cookies, which has led to the development of laws not just in the United States but also in the European Union.

Before keeping any cookies on a user's device that aren't absolutely essential, all websites that are focused on the European Union are required by the General Data Protection Regulation (GDPR) to get "informed permission" from users. This consent must be given voluntarily. The purpose of this criterion is to preserve the privacy of users and guarantee that they have control over the cookies that are kept on their devices by requiring them to comply with it.

5.3.2 What is TCF

The General Data Protection Regulation (GDPR) of the European Union went into force on May 25, 2018, and as a result, the landscape of data privacy in Europe has undergone a substantial transformation. Its primary objective was to bring data privacy laws in Europe into greater conformity with one another, with the goals of providing individuals with a greater degree of control and transparency over their personal data and establishing a higher bar for organizations seeking to legally process personal information.

The IAB Europe Transparency and Consent Framework (TCF) is a one-of-a-kind solution for the GDPR consent requirement. It is the only commercial product that was developed by the industry for the industry, and it established a genuine industry-standard approach. When it comes to processing personal data or accessing and storing information on a user's device, such as cookies, advertising identifiers, device identifiers, and other tracking technologies, the primary goal of the TCF is to assist all parties in the digital advertising ecosystem in complying with the GDPR and the ePrivacy Directive. This is done in the context of personal processing data.

The TCF provides an environment in which website publishers are able to notify users about the data that is being collected, as well as the planned use of the data by their website and the partners of their website. The TCF makes it possible for businesses in the publishing and advertising sectors to speak with customers in a standardized manner about their agreement to receive relevant online advertising and content by providing a common vocabulary.

In this work we will analyze examples of websites using the TCF.

5.3.3 IAB Europe

The IAB Europe is a pan-European organization that acts as the voice of the community of people who are involved in digital marketing and advertising. The group seeks to give a united representation of the industry in political and public dialogue. Its membership is comprised of important media, technology, and marketing corporations as well as national IABs at the national level. IAB Europe is committed to achieving its purpose by working toward the establishment of frameworks, standards, and industry-wide initiatives that will encourage the expansion and success of enterprises that are active in the European market. The organization is committed to fostering cooperation among its member companies, with the goal of capitalizing on their combined capabilities to propel innovation within the digital marketing and advertising industry and mold its trajectory into the future.

Some of the IAB Members are:

Adobe, Springer, BBC, Bloomblerg, CNN, Ebay, Alphabet (Google), Huawei, Microsoft, Meta (Facebook - Whatsapp), Oracle, Yahoo.

5.4 Cookie scenarios

5.4.1 BBC.com scenario

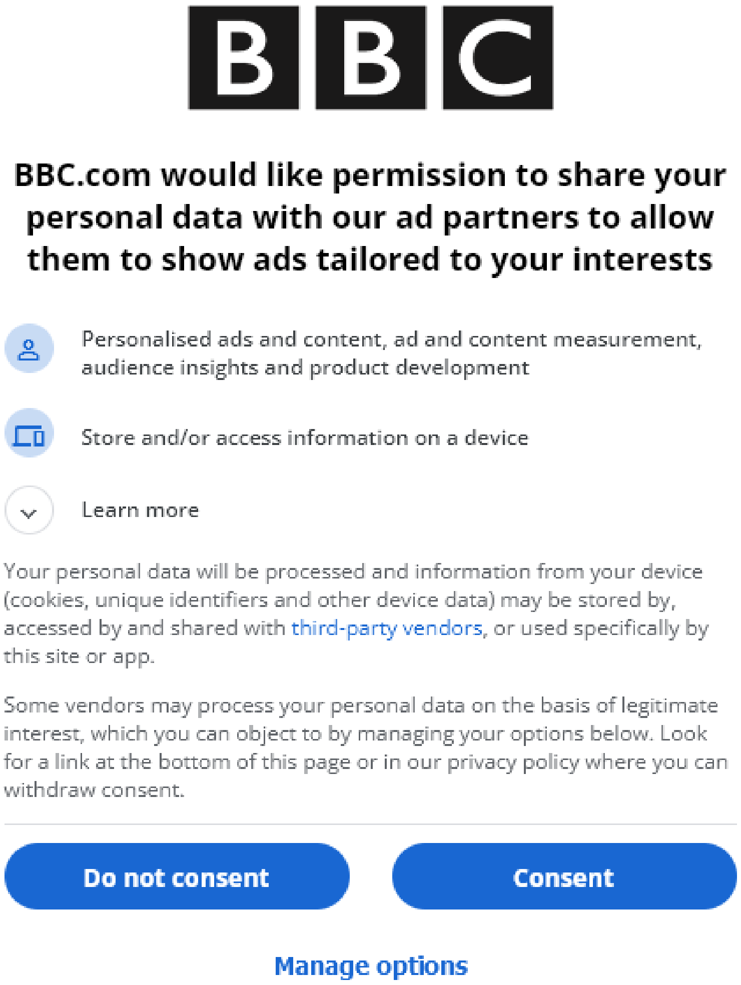
Alice wants to catch up on the latest international news about the current geopolitical situation, and she accesses the BBC.com web site.

A popup informs her that the site BBC.com would like permission to use and share her personal data with ad partners to allow them to show ads tailored to her interests.

She has three possible choices:

- Click "**Do not consent**" button
- Click "**Consent**" button
- Click "**Manage options**" button

We will evaluate the effects and the consequences of the scenario, in general and considering the three possible cases.



BBC

BBC.com would like permission to share your personal data with our ad partners to allow them to show ads tailored to your interests

Personalised ads and content, ad and content measurement, audience insights and product development

Store and/or access information on a device

Learn more

Your personal data will be processed and information from your device (cookies, unique identifiers and other device data) may be stored by, accessed by and shared with third-party vendors, or used specifically by this site or app.

Some vendors may process your personal data on the basis of legitimate interest, which you can object to by managing your options below. Look for a link at the bottom of this page or in our privacy policy where you can withdraw consent.

Figure 5.1: BBC.com cookies popup

5.4.1.a Case 1: "Do not consent"

Alice is privacy-aware and prefers to not have a personalized experience if this means to share personal information. She clicks "**Do not consent**".

EFFECTS:

- Her web experience is not tailored to her interest and the ads she receives are annoying and completely useless for her.
- She is not a native English speaker but she is presented with the

home page in English.

- The news she sees is not very relevant to her country.
- The weather forecast present in the BBC.com page is useless to her because is referred to another city.
- Although in this session her choice will be remembered by the system, in the next session she will again be presented with the same popup and asked to make her choice again.

5.4.1.b Case 2: "Consent"

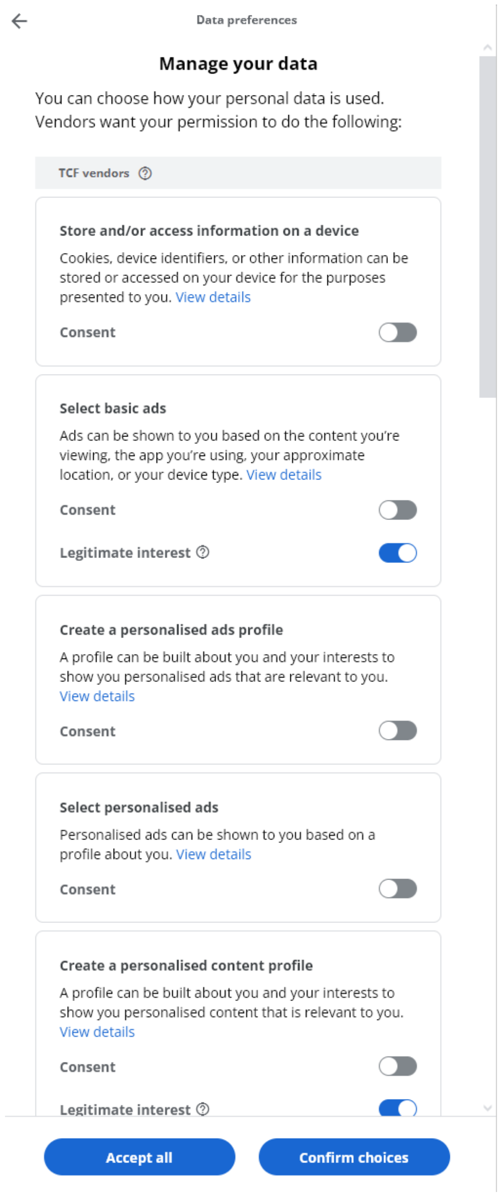
Alice is interested in personalizing her web experience and clicks "**Consent**".

EFFECTS:

- Her web experience is tailored to her interest and the ads she receives are useful to her.
- She is not a English native speaker ad she is presented with the home page in her language.
- The news she sees is relevant to her country.
- The weather forecast present in the BBC.com page refers to her city.
- Her choice will be remembered by the system and she will not be presented with the popup again until the relevant cookie expires.

5.4.1.c Case 3: "Manage options"

Alice is privacy-aware but she is also interested in personalizing her web experience, so she clicks "**Manage options**".



Manage your data

You can choose how your personal data is used. Vendors want your permission to do the following:

TCF vendors ⓘ

Store and/or access information on a device

Cookies, device identifiers, or other information can be stored or accessed on your device for the purposes presented to you. View details

Consent ○

Select basic ads

Ads can be shown to you based on the content you're viewing, the app you're using, your approximate location, or your device type. View details

Consent ○

Legitimate interest ⓘ ●

Create a personalised ads profile

A profile can be built about you and your interests to show you personalised ads that are relevant to you. View details

Consent ○

Select personalised ads

Personalised ads can be shown to you based on a profile about you. View details

Consent ○

Create a personalised content profile

A profile can be built about you and your interests to show you personalised content that is relevant to you. View details

Consent ○

Legitimate interest ⓘ ●

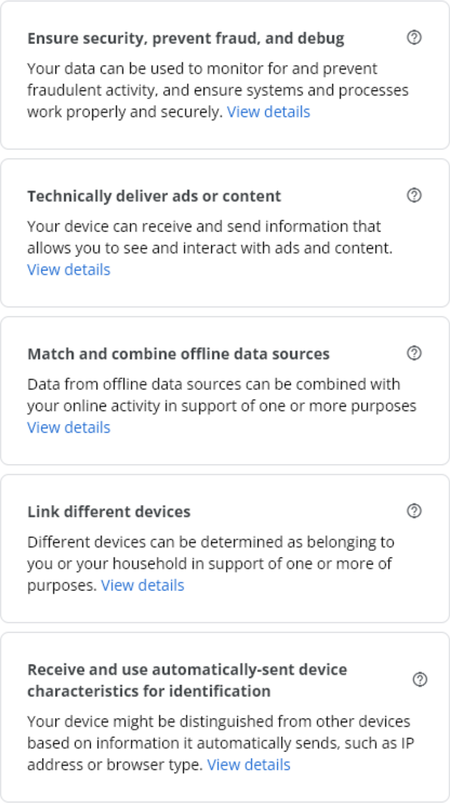
Figure 5.2: BBC.com cookies options popup

- She is presented with list of the of possible options to manage what information she wants or doesn't want to share.

EFFECTS:

- She can switch on and off the "**Consent**" option for some of the available choices.

- She can switch on and off the "**Legitimate interest**" option for some of the available choices.
- She is presented with a list of informative boxes with mandatory actions that the site will do to provide the service.



Ensure security, prevent fraud, and debug

Your data can be used to monitor for and prevent fraudulent activity, and ensure systems and processes work properly and securely. View details

Technically deliver ads or content

Your device can receive and send information that allows you to see and interact with ads and content. View details

Match and combine offline data sources

Data from offline data sources can be combined with your online activity in support of one or more purposes View details

Link different devices

Different devices can be determined as belonging to you or your household in support of one or more of purposes. View details

Receive and use automatically-sent device characteristics for identification

Your device might be distinguished from other devices based on information it automatically sends, such as IP address or browser type. View details

Figure 5.3: BBC.com cookies information popup

- She has the option to click the "**Accept all**" button to switch on all the proposed option or click "**Confirm choices**" to confirm the actual setup of the preferences.

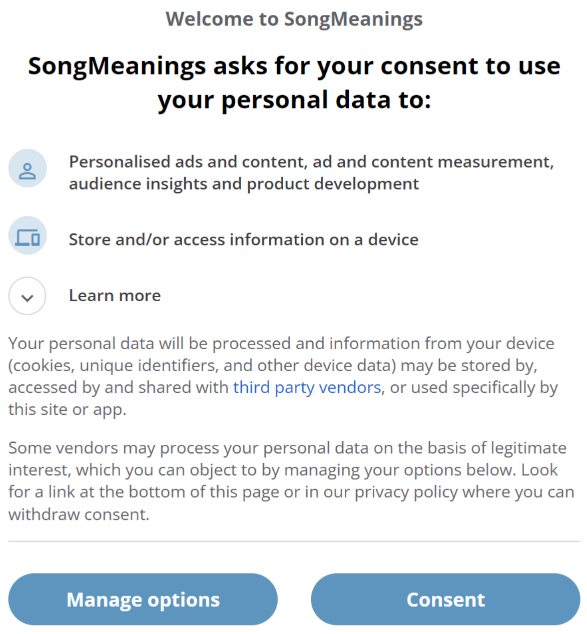
CASE 4: Alice does not agree with any of the options and decide to close the website.

EFFECTS:

- Alice can't receive any of the service proposed on the BBC.com website.
- She already disclosed some of the mandatory information she was informed about in case 3 (e.g., IP address, user-agent string).
- She already disclosed some of the information that falls within the "legitimate interest"

5.4.2 songmeanings.com scenario

Bob wants to search for the lyrics of a song and read opinions on its meaning. He opens the songmeanings.com web site. A popup informs him that the site songmeanings.com would like permission to use and share his personal data with ad partners to allow them to show ads tailored to his interests.



Welcome to SongMeanings

SongMeanings asks for your consent to use your personal data to:

Personalised ads and content, ad and content measurement, audience insights and product development

Store and/or access information on a device

Learn more

Your personal data will be processed and information from your device (cookies, unique identifiers, and other device data) may be stored by, accessed by and shared with third party vendors, or used specifically by this site or app.

Some vendors may process your personal data on the basis of legitimate interest, which you can object to by managing your options below. Look for a link at the bottom of this page or in our privacy policy where you can withdraw consent.

Figure 5.4: songmeanings.com cookies popup

We can notice the **do not consent** option is not enabled, this means that the Bob *must* manage the options if he wants to disable the sharing of some information or disable some cookie features while he is visiting the website. Even if the GDPR prohibits to the web site owner to making access conditional to the acceptance of the cookies, to disable the "do not consent" option means to slow down the process to the point the user is somewhat forced to click the "consent" button to use the site.

5.5 Actual mitigating solutions

5.5.1 Web browsers extensions

The capabilities of a web browser may be expanded by installing add-ons known as extensions. These add-ons are created utilizing common web-based programming languages such as HTML, CSS, and JavaScript. In addition, extensions not only have access to their very own collection of JavaScript application programming interfaces (APIs), but they also have the ability to use any web APIs that are present on a web page as long as the website supports JavaScript. While compared to the constraints that they may encounter when employing code inside a web page, this provides developers with a larger variety of capabilities when working within an extension. The following is a short list of some of the things that can be done with extensions:

- Installing an add-on on a user's browser that brings additional features or information from a website into the user's browser may be a useful way to enhance or expand the functionality of a website. Make it easy for consumers to get information from the sites they visit so that you can enhance the quality of the service that you provide.

- The developer is able to assist users in blocking annoying adverts from appearing on websites, providing access to a trip guide anytime a nation or place is mentioned on a web page or reformatting page content in order to create a more uniform experience while reading online information. They are able to accomplish any of these things by modifying the information on websites or eliminating it entirely. Extensions provide users access to the HTML and CSS of a website as well as the ability to make modifications to those programming languages. This enables users to personalize the way the web is shown to them.

- Tools and new browsing options may be introduced. For instance, a taskboard may have additional features added to it, or QR code graphics can be produced from URLs, hyperlinks, or the text on a website. Both of these examples are possible. You may easily add new functionality to a browser by using the robust APIs that are made available by WebExtensions and the many user interface options that are available. Additionally, you have the ability to upgrade the features or functionality of almost any website; it does not even have to be your own website for this to be possible.

- Add development tools: the developer may extend the functionality of the developer tools that are pre-installed with the browser by adding a new tab to the developer toolbar. In this example, the developer tools come pre-installed with Google Chrome. [49]

Using extensions for the purpose of automatically managing cookies has become increasingly common, and to exemplify this point, we can cite two widely used and highly regarded cookie management extensions that have garnered considerable attention in the field.

Descriptions of the extensions from the Chrome Web Store:

- **I don't care about cookies** (800.000+ users)
  «Remove annoying cookie warnings from almost all 'infected' websites! [...] The EU regulations require that any website using cookies must get user's permission before installing them. Imagine how irritating that becomes when you surf anonymously or if you delete cookies automatically every time you close the browser.» [101]

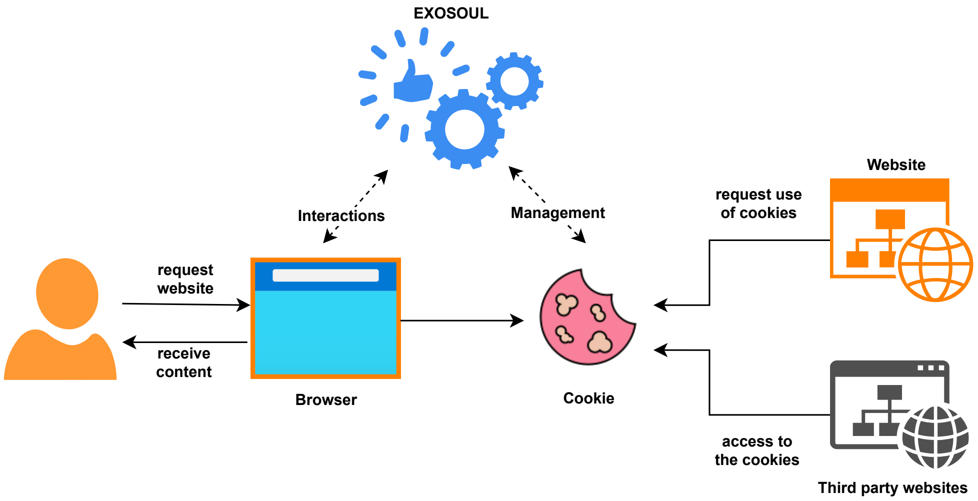


Figure 5.5: Diagram of the cookies workflow managed by EXOSOUL.

- **Cookie-Editor** (800.000+ users)
  «Cookie-Editor is designed to have a simple to use interface that let you do most standard cookie operations. It is ideal for developing and testing web pages or even manual management of cookies for your privacy.» [50]

5.6 Make use of EXOSOUL in cookies management

The EXOSOUL software exoskeleton is designed to balance power between users and systems in the digital world, empowering the user with the ability to reflect her ethical profile in digital interactions [22]. Once properly trained [15], EXOSOUL will be able to manage the interactions with websites, managing the cookies for the requested website and the future access to the cookies from third party websites. It is evident how the use of a customized software exoskeleton is a huge advantage over off-the-shelf commercial solutions (like Apple Intelligent Tracking Prevention) because it uses the user's own profile to dynamically manage the interaction and because the user remains the sole party in charge of managing her or his own data, without the burden of having to materially worry about continually keeping up with new profiling technologies and related mitigation solutions.

Because EXOSOUL takes advantage of hybrid profiling methodologies, involving the use of model-driven pre-trained profiles other than fine-grained personalized profile of the user that is automatically trained observing users behaviour, we can outline how web browsing mediation could be applied based on privacy profiles emerged from preliminary EXOSOUL model-driven categories developing [59]. In [59] we can find three main privacy categories (mainly based on the Westin's segmentation) and the further elaboration of the center category in two refined sub-categories:

- INATTENTIVE
- INVOLVED/ATTENTIVE
- SOLICITOUS

At this stage, it is essential to assign particular preferences to each unique profile, and there are a few different ways that this may be accomplished. Since we are conducting our analysis of the preliminary assignment of a particular set of preferences to a profile in a top-down manner, the assignment will consist of a set of default preferences. These preferences can, for instance, be selected by an experienced analyst, or they can be collected through the use of a unique questionnaire that can relate the ethical and privacy profile of the user to the particular preferences that will be assigned by default. In either case, the assignment will involve a default set of preferences. Users will have a profile with initial preferences set by default. The profile can be customized based on the user's own preferences, which can be gathered through websites they visit or through direct management of preferences.

In this example, we assign to each of these categories a certain set of preferences for cookies management:

- INATTENTIVE
  - Consent all cookies: YES

  - Manage cookie options: NO

- INVOLVED
  - Consent all cookies: NO
  - Manage cookie options: YES
    - Store and/or access information on a device: YES
    - Select basic ads: YES
    - Create personalised ads profile: NO
    - Create personalized content profile: NO
    - Select personalised content: NO
    - Measure ad performance: YES
    - Measure content performance: YES
    - Apply market research to generate audience insights: YES
    - Develop and improve products: YES
    - Ensure security, prevent fraud, and debug: YES
    - Technically deliver ads or content: YES
    - Match and combine offline data sources: NO
    - Link different devices: YES
    - Receive and use automatically-sent device characteristics for identification: YES
    - Use precise geolocation data: NO
- ATTENTIVE
  - Consent all cookies: NO
  - Manage cookie options: YES
    - Store and/or access information on a device: YES
    - Select basic ads: YES
    - Create personalised ads profile: NO
    - Create personalized content profile: NO
    - Select personalised content: NO
    - Measure ad performance: NO

        - Measure content performance: NO
        - Apply market research to generate audience insights: YES
        - Develop and improve products: YES
        - Ensure security, prevent fraud, and debug: NO
        - Technically deliver ads or content: NO
        - Match and combine offline data sources: NO
        - Link different devices: NO
        - Receive and use automatically-sent device characteristics for identification: NO
        - Use precise geolocation data: NO

- SOLICITOUS
    - Consent all cookies: NO
    - Manage cookie options: NO

5.7 How EXOSOUL can empower the user managing cookies

The advent of the Internet has brought in a wide range of options and possible threats for the security of personal data and the protection of user's privacy. With the introduction of the EXOSOUL Software Exoskeleton, a new level of security and privacy protection could be established, enabling users to manage cookies in an automatic and safe manner. This system provides users with a method to safeguard their data and personal information from being compromised, while letting them to fully use the web's benefits. A software exoskeleton that handles cookies automatically by using a user's personal profile might be incredibly valuable for a wide range of reasons. First, it may assist users in maintaining their privacy and security by enabling them to manage the types of cookies that websites can keep on their machines. Users may opt to store just particular kinds of cookies or to remove all cookies at the conclusion of each session. This would shield users from possible security threats associated with tracking data and rogue websites. Second, the software exoskeleton may assist users save time by facilitating rapid access to previously visited websites. By saving cookies for each website, the program may immediately identify the user and give them with material tailored to their prior visits. This may significantly reduce the amount of time people spend exploring web pages. Thirdly, a software exoskeleton may help users save money by delivering customized adverts and offers based on their prior internet visits. The software exoskeleton may further protect users against fraudulent websites and monitoring data. By analyzing the cookies saved on the user's computer, the program is able to identify possible security threats and prevent access to particular websites. This may assist safeguard users against possible security dangers and fraudulent websites. Overall, a software system that automatically manages cookies by using a user's personal profile may be incredibly beneficial for a variety of reasons. It may save customers time and money, as well as safeguard their privacy and security. The software exoskeleton may give users with a personalized and safe surfing experience by letting them to manage the sorts of cookies saved on their computers. Users are often forced to accept cookies to access websites that monitor user behavior, and they may also be used to store personal data. This information may then be used for targeted advertising or by third parties to obtain a user's personal information. EXOSOUL enables users to handle cookies in a more secure and efficient manner. This software application can identify and prevent cookies that may be used to follow a user's activity or obtain access to their personal data. Additionally, the software exoskeleton may be used to erase cookies automatically when they are no longer required. This helps safeguard a user's privacy, since they will no longer be required to erase cookies manually. In addition to preserving a user's privacy, EXOSOUL enables users to use the web's capabilities. Users are able to tailor the websites they visit and the material they view taking advantage of the user's personal profile. This enables consumers to customise their online experience to their own requirements and interests, and may be an useful method to guarantee that only relevant and engaging material is provided. EXOSOUL may also be utilized to make the browsing experience more secure. The software exoskeleton may guarantee that a user's personal information is not being monitored or shared with third parties by automatically managing cookies. This may help to lessen the chance of a user's data being compromised and secure the user's identity and privacy-sensible data. By managing cookies automatically, users may surf the web more safely and protect their personal information. In addition, taking into account the user's personal profile, individuals may tailor their online experience to their own requirements and interests.

5.8 Architecture of the EXOSOUL cookies management browser extension

Based on the Chrome Developers' Extensions Developing Guidelines [58], the architecture of a cookies management browser extension based on the EXOSOUL user's personal privacy profile will consist of the following components:

- **User interface**: A user-friendly interface is needed to display the privacy profile and allow the user to manage their settings. This could be in the form of a popup window or a toolbar button.
- **Content script**: A content script runs in the context of the web page and has access to its Document Object Model (DOM)[45] and cookies. It communicates with the background script to retrieve and update cookie information.
- **Storage and Data Persistence**: The cookie information must be stored in a persistent storage area, such as local storage or indexedDB[46] so that the information can be retained between browser sessions.
- **Request interception and background processing**: The background script listens for network requests and modifies them as necessary to add, delete, or modify cookies based on the user's privacy profile. This can be done using browser APIs such as the Web Request API[47] or the WebNavigation API[48].

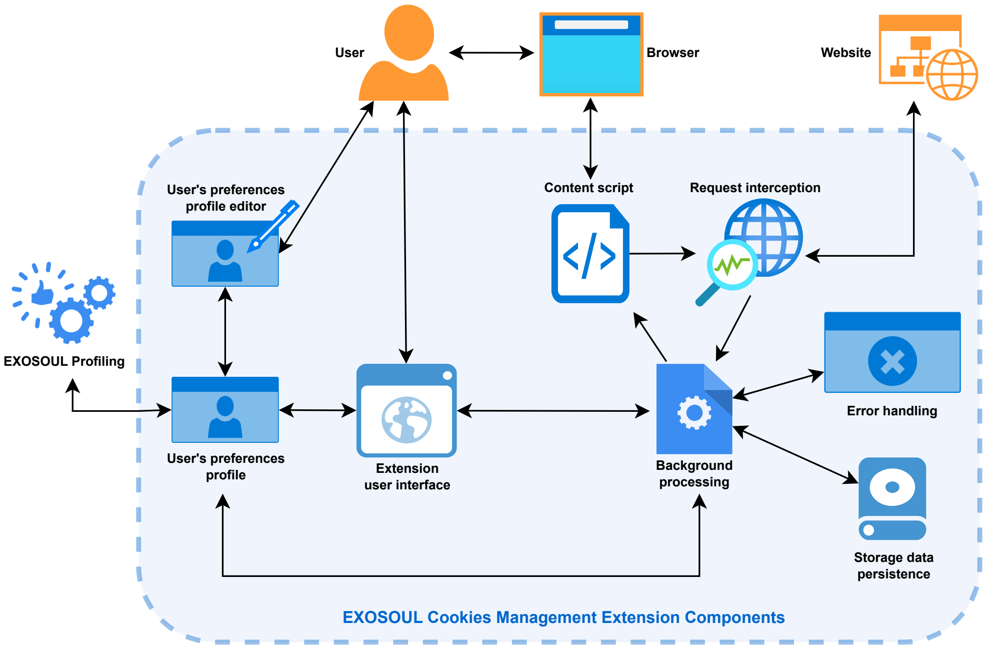


Figure 5.6: Diagram of the architecture of the EXOSOUL cookies management browser extension.

- **Privacy profile processing**: The background script must process the user's privacy profile to determine the cookies management for the current website. For example, if the user's preferences profile specifies that they never want to allow cookies from a certain type of website, the background script can block those cookies.
- **User interaction**: The user interface must allow the user to view and manage their privacy profile, including adding, deleting, or modifying privacy settings for specific domains.
- **Error handling**: The extension must include error handling and logging to diagnose and resolve any issues that may arise while managing cookies based on the user's privacy profile.

In Figure 5.6 shows the architecture of the EXOSOUL web browser extension with all its components.

Chapter 6

Conclusion

The purpose of this chapter is to present a review of our work's contributions to the field, as well as its limitations and future directions. We first present an assessment of our research's contributions and highlight their relevance in advancing the state of the art, then we discuss the limits of our study and offer suggestions for future research that may permit to overcome these limitations.

6.1 Summary of contributions

This section will explore in detail all the contributions of this dissertation with respect to the proposed chapters.

Some of the results and contributions obtained from this research have been published or are in the process of being published in the following articles:

- Patrizio Migliarini, Gian Luca Scoccia, Marco Autili, Paola Inverardi, 2020. On the elicitation of privacy and ethics preferences of mobile users. In Proceedings of the IEEE/ACM 7th International Conference on Mobile Software Engineering and Systems (pp. 132-136).
- Costanza Alfieri, Paola Inverardi, Patrizio Migliarini, and Massimiliano Palmiero. Exosoul: ethical profiling in the digital world.

HHAI2022 - Volume 354 of Frontiers in Artificial Intelligence and Applications, IOS Press 10.3233/FAIA220194, ISBN print: 978-1-64368-308-9, ISBN online: 978-1-64368-309-6, 2022.

- Davide Di Ruscio, Paola Inverardi, Patrizio Migliarini, and Phuong T Nguyen. Leveraging privacy profiles to empower users in the digital society. arXiv preprint arXiv:2204.00011, 2022 - submitted to journal.

- Paola Inverardi, Patrizio Migliarini, and Massimiliano Palmiero. Systematic Review on Privacy Categorization. 2023 - submitted to journal.

6.1.1 Systematic review contributions

Twenty-four papers were discovered in the systematic review, indicating limited investigations on privacy classification. The studies investigated nine primary contexts/domains: general, economy, mobile applications, health, social networks, computing, location sharing, e-commerce, and the internet. Most of the studies focused on general and economic concerns, and most of Westin's efforts were related to economics and marketing. Recent research indicates that ethical profiles, attitudes, beliefs, and moral ideals influence privacy-related behavior as much as domain/-context. The amount of sensitivity of the categories, such as health and finances versus social media and GPS-enabled apps, may influence individuals' privacy-related attitudes and behaviors. According to Nissenbaum, privacy can have several meanings depending on the context or domain. Depending on the domain or situation, subjects may be more or less circumspect about exposing personal information. In order to effectively reflect privacy preferences in terms of individual attitudes, future research should employ a more general domain approach. Instead of contextual variables or distinctive attitudes and practices in a certain domain/context, an individual's ethical profile governs privacy preferences; this implies that to classify users correctly, it is necessary to consider their stable attitudes toward ethics and privacy. The literature review on privacy categorization revealed that various methodologies and instruments (self-report questionnaires, interviews, and data-set analyses) are utilized, with the questionnaire being the most prevalent instrument. Some studies use hybrid approaches, which can lead to more specific and detailed categorizations than model-driven approaches. A pure data-driven approach has limitations and biases, whereas a self-report questionnaire or interview is limited by the privacy paradox. Refining privacy categories necessitates multidisciplinary and hybrid approaches, including perspectives from the human sciences. Only a combination of computer science and human/social science is capable of analyzing the complexity of privacy behavior. Digitization and technological advances have significantly impacted the development of categorization research. Recent research has demonstrated that technological advancement has played a defining role in categorizing user preferences and ethics. The widespread adoption of digital technologies has increased the user's digital fingerprint, allowing researchers to identify patterns and make informed decisions. This has led to the development of more refined privacy protection solutions for users. However, some studies have expanded the classification to include ethical and philosophical considerations in addition to technical behaviors and skills. The classification of user privacy behavior has evolved from Westin's original three segments (Unconcerned, Pragmatic, and Fundamentalist) to more refined and detailed classifications, with some studies proposing as many as seven distinct profiles. Except for Hoofnagle et al. (2014), which is a theoretical analysis of Westin's work, all of these studies are concerned with user interactions in the digital world.

6.1.2 Eliciting user's privacy profile contributions

In chapter 3, we evaluated the consistency of a user dataset's self-assessment of privacy categories. We divided the users into four privacy categories: Privacy Conservative (Class 0), Unconcerned, Fence-Sitting, and Advanced User (Class 3). We used a neural network classifier trained on the privacy settings of the users and their self-assigned labels to evaluate the consistency. We then evaluated the performance of the classifier using 10-fold cross-validation. Results indicated that prediction performance was poor, with ROC curves sloping in the direction of a random guess and AUC values below 0.65. This indicated that the training data contained negative results and noise. A significant proportion (96.20%) of the users had very similar users in different self-assessed categories, according to additional analysis. To address these findings, experiments were conducted on various subsets of a questionnaire to determine the optimal set of questions for user clustering. Using subset A as input yielded the most compact clusters, while privacy-relevant questions yielded the most discriminatory clusters. On this basis, we proposed a more practical method for assigning new users to clusters that avoid repetitive clustering. The final performance was evaluated using ROC curves, which revealed that our proposed method outperformed conventional clustering. In conclusion, our findings indicate that users' self-assessments of their privacy attitudes do not always correspond to their actual behavior, leading us to conclude that there is a need for a more effective method of understanding and classifying privacy preferences. Our second contribution can offer a solution to the problem: the creation of PisaRec, a recommender system designed to provide users with privacy settings that reflect their individual preferences. PisaRec reduces the burden of setting privacy configurations by providing users with recommendations tailored to their specific needs and preferences. This solution could be implemented in overarching systems like EXOSOUL to assist user's and improve their online experience.

6.1.3 Developing a user's ethical profile contributions

Chapter 4 aimed to develop ethical profiles capable of predicting digital activities. Two clustering solutions, k = 2 and k = 4, combining Ethics Position Theory (idealism and relativism), personality (honesty/humility, conscientiousness, Machiavellianism, and narcissism), and worldview (normativism), were used to develop the ethical profiles. Normativism was utilized as a criterion for determining whether an individual adheres to societal norms. The research revealed that the 2-cluster solution, which demonstrated a normativism-only overlap between the clusters, is implementable. The discriminant analysis revealed that a significant proportion of instances were correctly categorized. The "virtuous" cluster was more likely to comply with copyright and privacy laws than the "opportunist" cluster, indicating that the "virtuous" cluster adheres to moral principles. In contrast, the "opportunist" cluster is likelier to violate norms for personal gain. The "virtuous" cluster also received higher scores for privacy settings and information registration on the internet, indicating that they are more vigilant and circumspect. In terms of Machiavellianism and narcissism, there was overlap between the "legalist," "virtuous," and "sensible" clusters in the 4-cluster solution. The "opportunist" cluster had higher scores for violations of privacy and copyright. The "virtuous" cluster was the only one with lower caution scores, indicating that ethical profiles can be composed of multiple components, including ethical personality traits. The 2-cluster solution could be a foundation for developing more refined ethical profiles, whereas the 4-cluster solution is still in development and requires additional testing. Overall, these findings demonstrate that ethical profiles can be described in positive and negative terms, each having repercussions on one's digital behavior.

6.1.4 EXOSOUL cookies management browser extension contributions

The proposal for a web browser extension aims to give the EXOSOUL exoskeleton a practical application. The objective is to provide the user with a personalized web browsing experience that is consistent with their personal ethical values, particularly concerning cookie management. This section proposes a comprehensive architecture for web browser extensions that will enable the development and implementation of EXOSOUL ethical profiling management for actual web navigation. By incorporating EXOSOUL technology into a web browser extension, we intend to improve the user's online experience. The extension will be tailored to the user's ethical profile, as determined by his or her individual preferences and values. The extension will focus primarily on managing cookies. The proposed extension will allow users to manage cookies according to their ethical profile, such as by setting preferences automatically or blocking or allowing specific types of cookies. In conclusion, the proposal for a web browser extension that incorporates the EXOSOUL exoskeleton presents a unique opportunity to offer users a more ethical and individualized web browsing experience. The interaction with the EXOSOUL ethical profiling and the active cookies management proposed by this extension will ensure that the user's web browsing experience is tailored to his or her ethical/privacy profile and preferences.

6.2 Limitations and further works

In the following, we discuss limitations and any further research that may be necessary.

6.2.1 From ethics to practical applications

A very recent work [181] addresses the problem of eliciting ethical requirements for autonomous systems that interact with users. It proposes a procedure for deriving practical rules from a collection of high-level normative principles in order to allow autonomous agents to pick and execute the most normatively advantageous action. Typically, the norms discussed in [181] are of a *social*, *legal*, *ethical*, *empathic*, or *cultural* (SLEEC) character and are abstractly stated. The procedure presented in [181], which transforms normative principles into SLEEC rules, entails defining normative principles, addressing SLEEC issues, identifying and resolving SLEEC disputes, and creating both preliminary and complicated normatively-relevant rules. In doing so, it bridges the gap between normative principles and operational practice, thereby guiding the development of autonomous agents and positioning them to be SLEEC-sensitive or SLEEC-compliant, which is particularly relevant given that these agents are frequently involved in tasks that directly and potentially negatively impact human well-being. SLEEC standards are abstract, high-level concepts that are difficult to translate to rules autonomous agents may obey. [181] describes a five-stage iterative approach for refining these ideas into rules that autonomous agents may obey, while meeting the needs of end users and other stakeholders. Consideration of the operational environment and the agent's design precedes the selection of relevant high-level principles, such as rights-based principles, cultural norms, and legal codes (stage 1). Then, these concepts are mapped to the agent's capabilities, and a first set of rules is constructed (stage 2). The agent then evaluates the points of influence inside the system-process that may have an undesirable effect on an SLEEC standard, as well as legal considerations and impact evaluations (stage 3). Conflicts are uncovered and addressed by evaluating trade-offs or producing compromises (stage 4). Finally, via a process of iteration (stage 5), the rules are modified and new opponents are promoted. When the refining of rules has been exhausted or there is no longer confidence that a subsequent iteration would meaningfully refine the rules further, the iteration process is complete. This will allow the system to make informed decisions and take the action that is deemed to be the most ethically and morally beneficial for the user. Utilizing this analysis technique will result in a robust and dependable method for guiding autonomous agents toward ethically preferable actions, thereby enhancing their overall decision-making abilities. The ethical requirements of an autonomous system may be defined with the use of methods like the one described. Those requirements, once explicit, could be useful to check compatibility with the user's ethics when her or him interacts with the system in approaches like EXOSOUL.

6.2.2 Privacy categorization further perspectives

Due to the exponential increase in personal information generated, stored, and shared online and due to technological advancements, privacy has become a critical concern in our increasingly interconnected society [54, 56, 202]. The need for privacy protection has led to the development of privacy profiling and categorization, which aim to comprehend and predict the privacy behavior of individuals [11, 173, 33]. Future research in this field should focus on developing more detailed and precise privacy categories. Privacy categories are determined mainly by demographic information and technical device characteristics. However, these categories do not always account for the psychological and cultural factors influencing an individual's privacy preferences and behavior [87, 104]. Future research should investigate multidisciplinary approaches, drawing on insights from psychology, sociology, philosophy, and other relevant fields, to develop a more nuanced and prescriptive understanding of privacy. For instance, the concept of personae can be used to comprehend privacy preferences and behavior better. Similarly, investigating privacy philosophies, such as those that emphasize individual control over personal information, can provide valuable insight into a person's privacy beliefs and preferences. The ultimate objective of this effort should be to develop privacy categories that correspond more closely to users' general privacy beliefs and practices. This will help predict and comprehend privacy preferences better, resulting in a more effective privacy protection framework. In addition, this work should be conducted transparently and ethically, taking into account the privacy rights of individuals and ensuring that any collected data is used solely for research purposes and anonymized appropriately to protect personal information. In conclusion, privacy profiling and categorization are essential fields that will continue to play a crucial role in protecting personal data in the digital age. A future research study in this area should define privacy categories through various multidisciplinary approaches and perspectives, pursuing more predictive categories such as personae and philosophies. This will assist in comprehending and protecting the privacy rights of individuals in an increasingly interconnected world.

6.2.3 Eliciting user's privacy profile advancements areas

Each of the following area has the potential to improve our findings and contribute to the advancement of the results that are presented in Chapter 3.

- Size of the Data: Despite the fact that the dataset we used in our experiments is relatively recent and accurately reflects the privacy practices of modern users, its size is relatively small. Experiments with a larger dataset that considers a wider variety of ethical considerations in addition to privacy are being planned as part of the next phases of the project.

- Bias: While we have endeavored to mitigate potential bias in autonomously generating user profiles and grouping the data, there may be limitations to our findings. To increase the reliability of our results, we performed a semantic analysis on the clusters generated by our proposed method and confirmed their precision. However, additional research is necessary to resolve any remaining limitations and identify potential confounding variables that may have influenced our findings. In future research endeavors, this will enhance the robustness and generalizability of our method.

- Generalizability: One of the most important factors determining the applicability and relevance of our findings in the real world is the extent to which they can be applied to similar situations in other contexts. In order to confirm the applicability of our findings in other contexts, we are planning additional experiments.

Advancements in these topics will contribute to the improvement and understanding of privacy profiles, enabling a better design of systems like EXOSOUL and their practical implications.

6.2.4 Developing user's ethical profile further works

The results presented in chapter 4 are encouraging and have opened up numerous areas for further research refinement and investigation. Future efforts will concentrate on enhancing our comprehension of the clusters identified as "legalist," "sensible," "virtuous," and "opportunistic." This will be accomplished by taking into account additional variables besides moral orientation and personality traits. Include scales that measure other ethical dimensions, such as integrity, authenticity, dogmatism, moral disengagement, sadism, and normlessness. Personality variables such as openness to experience and neuroticism could be added to the clustering process to provide additional insight. In addition, it would be intriguing to investigate the possibility of subdividing the "virtuous" and "opportunist" clusters into smaller, more distinct subclusters. The utilization of machine learning algorithms to the same research questions is another possibility for further research. The objective is to evaluate how well these algorithms predict the digital behavior of users based on their ethical stances and personality traits. It is essential to note that the present study's sample was restricted to university students and this can affect the generalizability of the results to the entire population. Future research should investigate the role of ethical profiles in predicting digital behavior in various age groups, educational levels, and cultural contexts. This will contribute to a deeper understanding of the relationships between ethics, personality, and digital behavior, as well as shed light on how these factors can be used to inform ethical decision-making in the digital domain.

6.2.5 Further applications of EXOSOUL in the web domain

Considering the multiplicity of weaknesses with respect to the security of information transmission and thus the possibility of unwanted profiling, and considering the high technical level required to understand some of these flaws, it would be desirable to develop an integrated solution that would rebalance the power between the user and the entire chain of systems involved, interfacing as closely as possible with the user, protecting and empowering her. Some of the web domain possible weaknesses in profiling-mitigation strategies includes IP and DNS logging, client/browser fingerprinting, etc. EXOSOUL can provide an extra layer of security to protect web applications and websites from potential threats, can provide users with an extra layer of protection against attacks such as profiling, data theft, and malicious code injection. This type of technology is becoming increasingly important in today's online world, as online users are becoming more and more vulnerable to data breaches, identity theft, and other cyber-crimes. EXOSOUL can be a valuable tool for enhancing web security against profiling. By providing an extra layer of security, it can help protect users from malicious code injections, data theft, profiling, phishing, and other malicious activities. In addition, it can also help protect users from certain types of network attacks. All of these benefits make a software exoskeleton a useful tool for protecting users from potential threats.

Bibliography

[1] Association for Computing Machinery U.S. Public Policy Council (USACM). Statement on Algorithmic Transparency and Accountability. . https://www.acm.org/binaries/content/assets/public-policy/2017usacmstatementalgorithms.pdf, 2018.

[2] European Group on Ethics in Science and New Technologies. statement on artificial intelligence, robotics and 'autonomous' systems. https://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf, 2018.

[3] Partnership on AI. https://www.partnershiponai.org, 2018.

[4] The European Commission's High-Level Expert Group on Artificial Intelligence, Draft Ethics Guidelines for Trustworthy Ai. https://ec.europa.eu/digital-single-market/en/news/nhave-your-say-european-expert-group-seeks-feedback-draft-nethics-guidelines-trustworthy, 2018.

[5] The IEEE Global hi Ethics of Autonomous and Intelligent Systems. https://standards.ieee.org/develop/indconn/ec/autonomous_systems.html, 2018.

[6] Why The Netherlands is the Globe's top Location for Self-driving Cars. https://www.consultancy.eu/news/1098/why-the-netherlands-is-the-globes-top-location-for-self-driving-cars, 2018.

[7] Gartner Top 10 Strategic Technology Trends for 2019. `https://gtnr.it/2CJJYGp`, 2019.

[8] The State of Autonomous Legislation in Europe. `https://autovistagroup.com/news-and-insights/state-autonomous-legislation-europe`, 2019.

[9] The 93rd United States Congress. Privacy act. 1974.

[10] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, and et al. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Comput. Surv.*, 50(3), August 2017.

[11] Alessandro Acquisti and Jens Grossklags. Privacy and rationality in individual decision making. *IEEE security & privacy*, 3(1):26–33, 2005.

[12] Charu Aggarwal. Neighborhood-based collaborative filtering. In *Recommender systems: The textbook*, pages 29–70. Springer International Publishing, Cham, 2016.

[13] Jamal A Al-Khatib, Mohammed I Al-Habib, Naima Bogari, and Najah Salamah. The ethical profile of global marketing negotiators. *Business Ethics: A European Review*, 25(2):172–186, 2016.

[14] Mark S Aldenderfer and Roger K Blashfield. Cluster analysis. newberry park, 1984.

[15] Costanza Alfieri, Paola Inverardi, Patrizio Migliarini, and Massimiliano Palmiero. Exosoul: ethical profiling in the digital world. *HHAI2022 - Volume 354 of Frontiers in Artificial Intelligence and Applications, IOS Press 10.3233/FAIA220194, ISBN print: 978-1-64368-308-9, ISBN online: 978-1-64368-309-6*, 2022.

[16] Anita Allen. Privacy and justice: Transnational values for the digital age (video). 2022.

[17] Gordon W Allport. Pattern and growth in personality. 1961.

[18] Gordon Willard Allport. Personality: A psychological interpretation. 1937.

[19] Ashwaq Alsoubai, Reza Ghaiumy Anaraky, Yao Li, Xinru Page, Bart Knijnenburg, and Pamela J Wisniewski. Permission vs. app limiters: Profiling smartphone users to understand differing strategies for mobile privacy management. In *CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2022.

[20] Michael C Ashton and Kibeom Lee. Empirical, theoretical, and practical advantages of the hexaco model of personality structure. *Personality and social psychology review*, 11(2):150–166, 2007.

[21] Michael C Ashton and Kibeom Lee. The hexaco–60: A short measure of the major dimensions of personality. *Journal of personality assessment*, 91(4):340–345, 2009.

[22] Marco Autili, Davide Di Ruscio, Paola Inverardi, Patrizio Pelliccione, and Massimo Tivoli. A software exoskeleton to protect and support citizen's ethics and privacy in the digital world. *IEEE Access*, 7:62011–62021, 2019.

[23] Naveen Farag Awad and Mayuram S Krishnan. The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS quarterly*, pages 13–28, 2006.

[24] Heenkenda Mudiyanselage Ruwan Jayawickrama Bandara, Mario Fernando, and Md Shahriar Akter. Is the privacy paradox a matter of psychological distance? an exploratory study of the privacy paradox from a construal level theory perspective. *Proceedings of the 51st Hawaii International Conference on System Sciences | 2018 (pp. 3678-3687). Hawaii, United States: University of Hawaii at Mano.*

[25] Stanley I Benn. Privacy, freedom, and respect for persons. In *Privacy & personality*, pages 1–26. Routledge, 2017.

[26] Bettina Berendt, Oliver Günther, and Sarah Spiekermann. Privacy in e-commerce: Stated preferences vs. actual behavior. *Communications of the ACM*, 48(4):101–106, 2005.

[27] Jaspreet Bhatia, Travis D Breaux, Joel R Reidenberg, and Thomas B Norton. A theory of vagueness and privacy risk perception. In *2016 IEEE 24th International Requirements Engineering Conference (RE)*, pages 26–35. IEEE, 2016.

[28] Christopher M. Bishop. *Neural networks for pattern recognition.* Oxford University Press, Inc., New York, NY, USA, 1995.

[29] Tom Buchanan, Carina Paine, Adam N Joinson, and Ulf-Dietrich Reips. Development of measures of online privacy concern and protection for use on the internet. *Journal of the American society for information science and technology*, 58(2):157–165, 2007.

[30] J. Peter Burgess, L. Floridi, A. Pols, and J. van den Hoven. Towards a digital ethics - edps ethics advisory group. `https://edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf`, 2018.

[31] J. A. Burns and G. M. Whitesides. Feed-forward neural networks in chemistry: Mathematical systems for classification and pattern recognition. *Chem. Rev.*, 93:2583–2601, 1993. 377.

[32] Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O'donoghue, and Matthew Rabin. Regulation for conservatives: Behavioral economics and the case for" asymmetric paternalism". *University of Pennsylvania law review*, 151(3):1211–1254, 2003.

[33] Abdelberi Chaabane, Gergely Acs, Mohamed Ali Kaafar, et al. You are what you like! information leakage through users' interests. In *Proceedings of the 19th annual network & distributed system security symposium (NDSS).* Citeseer, 2012.

[34] Richard Christie and Florence L Geis. *Studies in machiavellianism.* Academic Press, 2013.

[35] Agata Chudzicka-Czupała. Ethical ideology as a predictor of ethical decision making. *International Journal of Management and Bussiness, 4 (1)*, pages 28–41, 2013.

[36] Roger Clarke. The digital persona and its application to data surveillance. *The information society*, 10(2):77–92, 1994.

[37] Taya R Cohen and Lily Morse. Moral character: What it is and what it does. *Research in organizational behavior*, 34:43–61, 2014.

[38] Taya R Cohen, Abigail T Panter, Nazlı Turan, Lily Morse, and Yeonjeong Kim. Agreement and similarity in self-other perceptions of moral character. *Journal of Research in Personality*, 47(6):816–830, 2013.

[39] Taya R Cohen, Abigail T Panter, Nazlı Turan, Lily Morse, and Yeonjeong Kim. Moral character in the workplace. *Journal of personality and social psychology*, 107(5):943, 2014.

[40] Katherine L Collison, Colin E Vize, Joshua D Miller, and Donald R Lynam. Development and preliminary validation of a five factor model measure of machiavellianism. *Psychological Assessment*, 30(10):1401, 2018.

[41] Jessica Colnago, Yuanyuan Feng, Tharangini Palanivel, Sarah Pearman, Megan Ung, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. Informing the design of a personalized privacy assistant for the internet of things. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[42] European Commission. Proposal for a regulation of the european parliament and of the council - laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. *EUR-Lex*, 2021.

[43] Ramón Compañó and Wainer Lusoli. The policy maker's anguish: Regulating personal data behavior between paradoxes and dilem-

mas. In *Economics of information security and privacy*, pages 169–185. Springer, 2010.

[44] Sunny Consolvo, Ian E Smith, Tara Matthews, Anthony LaMarca, Jason Tabert, and Pauline Powledge. Location disclosure to social relations: why, when, & what people want to share. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 81–90, 2005.

[45] MDN contributors. Document object model. `https://developer.mozilla.org/en-US/docs/Web/API/Document_Object_Model`, 2022.

[46] MDN contributors. indexeddb. `https://developer.mozilla.org/en-US/docs/Web/API/indexedDB`, 2022.

[47] MDN contributors. Request api. `https://developer.mozilla.org/en-US/docs/Web/API/Request`, 2022.

[48] MDN contributors. webnavigation api. `https://developer.mozilla.org/en-US/docs/Mozilla/Add-ons/WebExtensions/API/webNavigation`, 2022.

[49] MDN contributors. What are extensions? `https://developer.mozilla.org/en-US/docs/Mozilla/Add-ons/WebExtensions/What-are-WebExtensions`, 2022.

[50] cookie editor.cgagnier.ca. cookie-editor. `https://chrome.google.com/webstore/detail/cookie-editor/`, 2022.

[51] Paul T Costa Jr and Robert R McCrae. Four ways five factors are basic. *Personality and individual differences*, 13(6):653–665, 1992.

[52] Paul T Costa Jr and Robert R McCrae. *Personality in adulthood: A five-factor theory perspective*. Routledge, 2013.

[53] Lorrie Faith Cranor. Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. *J. on Telecomm. & High Tech. L.*, 10:273, 2012.

[54] Natalia Criado and Jose M Such. Implicit contextual integrity in online social networks. *Information Sciences*, 325:48–69, 2015.

[55] Mary J Culnan and Pamela K Armstrong. Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization science*, 10(1):104–115, 1999.

[56] Jon P Daries, Justin Reich, Jim Waldo, Elise M Young, Jonathan Whittinghill, Andrew Dean Ho, Daniel Thomas Seaton, and Isaac Chuang. Privacy, anonymity, and big data in the social sciences. *Communications of the ACM*, 57(9):56–63, 2014.

[57] James R Detert, Linda Klebe Treviño, and Vicki L Sweitzer. Moral disengagement in ethical decision making: a study of antecedents and outcomes. *Journal of applied psychology*, 93(2):374, 2008.

[58] Chrome developers. Extension development guidelines. `https://developer.chrome.com/docs/extensions/mv3/getstarted/`, 2022.

[59] Davide Di Ruscio, Paola Inverardi, Patrizio Migliarini, and Phuong T Nguyen. Leveraging privacy profiles to empower users in the digital society. *arXiv preprint arXiv:2204.00011*, 2022.

[60] Janna Lynn Dupree, Richard Devries, Daniel M. Berry, and Edward Lank. Privacy personas: Clustering users via attitudes and behaviors toward security practices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 5228–5239, New York, NY, USA, 2016. Association for Computing Machinery.

[61] Janna-Lynn Weber Dupree, Edward Lank, and Daniel M Berry. A case study of using grounded analysis as a requirement engineering method: Identifying personas that specify privacy and security tool users. *Science of Computer Programming*, 152:1–37, 2018.

[62] EDPS. Leading by example, The EDPS Strategy 2015-2019. `https://edps.europa.eu/sites/edp/files/publication/15-07-30_strategy_2015_2019_update_en.pdf`, 2015.

[63] Dag Elgesem. Privacy, respect for persons, and risk. *Philosophical perspectives on computer-mediated communication*, pages 45–66, 1996.

[64] Robert A Emmons. Narcissism: theory and measurement. *Journal of personality and social psychology*, 52(1):11, 1987.

[65] Seda Erzi. Dark triad and schadenfreude: Mediating role of moral disengagement and relational aggression. *Personality and Individual Differences*, 157:109827, 2020.

[66] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, page 226–231. AAAI Press, 1996.

[67] J. Larus et al. When Computers Decide: European Recommendations on Machine-Learned Automated Decision Making. .

[68] Tatjana Evas. *A Common EU Approach to Liability Rules and Insurance for Connected and Autonomous Vehicles: European Added Value Assessment: Accompanying the European Parliament's legislative own-initiative report.* Number QA-04-18-027-EN-N. 2018.

[69] Tom Fawcett. An introduction to roc analysis. *Pattern Recogn. Lett.*, 27(8):861–874, June 2006.

[70] Luciano Floridi. Luciano floridi—commentary on the onlife manifesto. In *The onlife manifesto*, pages 21–23. Springer, Cham, 2015.

[71] Luciano Floridi. Soft ethics and the governance of the digital. *Philosophy & Technology*, 31(1):1–8, 2018.

[72] Luciano Floridi. The green and the blue: a new political ontology for a mature information society. *Available at SSRN 3831094*, 2020.

[73] Luciano Floridi and Josh Cowls. A unified framework of five principles for ai in society. *Machine learning and the city: Applications in architecture and urban design*, pages 535–545, 2022.

[74] Donelson R Forsyth. A taxonomy of ethical ideologies. *Journal of Personality and Social psychology*, 39(1):175, 1980.

[75] Donelson R Forsyth. *Making moral judgments: Psychological perspectives on morality, ethics, and decision-making*. Routledge, 2019.

[76] Donelson R Forsyth, George C Banks, Michael A McDaniel, et al. A meta-analysis of the dark triad and work behavior: a social exchange perspective. *Journal of applied psychology*, 97(3):557, 2012.

[77] Donelson R Forsyth, Ernest H O'boyle, and Michael A McDaniel. East meets west: A meta-analytic investigation of cultural variations in idealism and relativism. *Journal of Business Ethics*, 83(4):813–833, 2008.

[78] Margalit Fox. Alan f. westin, who transformed privacy debate before the web era, dies at 83. *The New York Times*, 2013.

[79] Francis Fukuyama, Barak Richman, and Ashish Goel. How to save democracy from technology: ending big tech's information monopoly. *Foreign Aff.*, 100:98, 2021.

[80] Adrian Furnham, Steven C Richards, and Delroy L Paulhus. The dark triad of personality: A 10 year review. *Social and personality psychology compass*, 7(3):199–216, 2013.

[81] Oscar H Gandy Jr. The role of theory in the policy process: a response to professor westin. *Toward an Information Bill of Rights and Responsibilities*, 1995.

[82] Oscar H Gandy Jr. Public opinion surveys and the formation of privacy policy. *Journal of social issues*, 59(2):283–299, 2003.

[83] Christina Geary, Evita March, and Rachel Grieve. Insta-identity: Dark personality traits as predictors of authentic self-presentation on instagram. *Telematics and Informatics*, 63:101669, 2021.

[84] Alan Goldman. *Destructive leaders and dysfunctional organizations: A therapeutic approach.* Cambridge University Press, 2009.

[85] Hyman Gross. Privacy and autonomy. In *Privacy & Personality*, pages 169–181. Routledge, 2017.

[86] Kevin Gurney. *An Introduction to Neural Networks.* Taylor and Francis, Inc., USA, 1997.

[87] Cory Hallam and Gianluca Zanella. Online self-disclosure: The privacy paradox explained as a temporally discounted balance between concerns and rewards. *Computers in Human Behavior*, 68:217–227, 2017.

[88] Louis Henkin. Privacy and autonomy. *Columbia Law Review*, 74:1410, 1974.

[89] Hielke Hijmans and Charles D Raab. Ethical dimensions of the gdpr. *Commentary on the General Data Protection Regulation, Cheltenham: Edward Elgar (2018, Forthcoming)*, 2018.

[90] Mireille Hildebrandt. Defining profiling: a new type of knowledge? In *Profiling the European citizen*, pages 17–45. Springer, 2008.

[91] Vanmala Hiranandani. Privacy and security in the digital age: contemporary challenges and future directions. *The International Journal of Human Rights*, 15(7):1091–1106, 2011.

[92] Soraj Hongladarom and Soraj Hongladarom. *Philosophical foundations of privacy.* Springer, 2016.

[93] Chris Jay Hoofnagle and Jennifer M Urban. Alan westin's privacy homo economicus. *Wake Forest L. Rev.*, 49:261, 2014.

[94] Paola Inverardi. The european perspective on responsible computing. *Communications of the ACM*, 62(4):64–64, 2019.

[95] Paola Inverardi. The challenge of human dignity in the era of autonomous systems. In *Perspectives on Digital Humanism*, pages 25–29. Springer, Cham, 2022.

[96] 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, 2004.

[97] Anna Jobin, Marcello Ienca, and Effy Vayena. The global landscape of ai ethics guidelines. *Nature Machine Intelligence*, 1(9):389–399, 2019.

[98] Leonard Kaufman and Peter J. Rousseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis.* John Wiley, 1990.

[99] Tariq Iqbal Khan, Aisha Akbar, Farooq Ahmed Jam, and Muhammad Mohtsham Saeed. A time-lagged study of the relationship between big five personality and ethical ideology. *Ethics & Behavior*, 26(6):488–506, 2016.

[100] Kagan Kircaburun, Zsolt Demetrovics, and Şule Betül Tosuntaş. Analyzing the links between problematic social media use, dark triad traits, and self-esteem. *International Journal of Mental Health and Addiction*, 17(6):1496–1507, 2019.

[101] Daniel Kladnik. I don't care about cookies. `https://chrome.google.com/webstore/detail/i-dont-care-about-cookies/`, 2022.

[102] Bart Piet Knijnenburg and Alfred Kobsa. Increasing sharing tendency without reducing satisfaction: Finding the best privacy-settings user interface for social networks. In *ICIS*, 2014.

[103] Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th international joint conference on artificial intelligence - volume 2*, IJCAI'95, pages 1137–1143, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc.

[104] Spyros Kokolakis. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security*, 64:122–134, 2017.

[105] Ilir Kola, Myrthe L Tielman, Catholijn M Jonker, and M Riemsdijk. Predicting the priority of social situations for personal assistant agents. In *International Conference on Principles and Practice of Multi-Agent Systems*, pages 231–247. Springer, 2020.

[106] Ponnurangam Kumaraguru and Lorrie Faith Cranor. Privacy Indexes: A Survey of Westin's Studies. Technical Report CMU-ISRI-5-138, Institute for Software Research International, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, Dezember 2005.

[107] Mark R Leary, Paul D Knight, and Byron D Barnes. Ethical ideologies of the machiavellian. *Personality and Social Psychology Bulletin*, 12(1):75–80, 1986.

[108] Hosub Lee and Alfred Kobsa. Understanding user privacy in internet of things environments. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, pages 407–412, 2016.

[109] Kibeom Lee and Michael C Ashton. Psychometric properties of the hexaco personality inventory. *Multivariate behavioral research*, 39(2):329–358, 2004.

[110] Kibeom Lee and Michael C Ashton. *The H factor of personality: Why some people are manipulative, self-entitled, materialistic, and exploitive—and why it matters for everyone.* Wilfrid Laurier Univ. Press, 2013.

[111] Jialiu Lin, Bin Liu, Norman Sadeh, and Jason I Hong. Modeling users' mobile app privacy preferences: Restoring usability in a sea of permission settings. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*, pages 199–212, 2014.

[112] Heather Richter Lipford, Andrew Besmer, and Jason Watson. Understanding privacy settings in facebook with an audience view. *UPSEC*, 8:1–8, 2008.

[113] Bin Liu. Can machine learning help people configure their mobile app privacy settings? `https://kilthub.cmu.edu/articles/thesis/Can_Machine_Learning_Help_People_Configure_Their_Mobile_App_Privacy_Settings_/11591340`, 2020.

[114] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Zhang, Norman Sadeh, Alessandro Acquisti, and Yuvraj Agarwal. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Proceedings of the Twelfth USENIX Conference on Usable Privacy and Security*, SOUPS '16, pages 27–41. USENIX Association, 2016. event-place: Denver, CO, USA.

[115] Bin Liu, Jialiu Lin, and Norman Sadeh. Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help? In *Proceedings of the 23rd international conference on World wide web*, pages 201–212, 2014.

[116] Xiaoli Liu, Satu Tamminen, Xiang Su, Pekka Siirtola, Juha Röning, Jukka Riekki, Jussi Kiljander, and Juha-Pekka Soininen. Enhancing veracity of iot generated big data in decision making. In *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 149–154. IEEE, 2018.

[117] Yanchi Liu, Zhongmou Li, Hui Xiong, Xuedong Gao, and Junjie Wu. Understanding of internal clustering validation measures. In *Proceedings of the 2010 IEEE International Conference on Data Mining*, ICDM '10, page 911–916, USA, 2010. IEEE Computer Society.

[118] Mary Madden. Privacy management on social media sites. *Pew Internet Report*, 24:1–20, 2012.

[119] T Soni Madhulatha. An overview on clustering methods. *arXiv preprint arXiv:1205.1117*, 2012.

[120] Naresh K Malhotra, Sung S Kim, and James Agarwal. Internet users' information privacy concerns (iuipc): The construct, the scale, and a causal model. *Information systems research*, 15(4):336–355, 2004.

[121] Patricia Yancey Martin and Barry A Turner. Grounded theory and organizational research. *The journal of applied behavioral science*, 22(2):141–157, 1986.

[122] Robert R McCrae and Paul T Costa. Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, 52(1):81, 1987.

[123] Aleecia M McDonald and Lorrie Faith Cranor. The cost of reading privacy policies. *Isjlp*, 4:543, 2008.

[124] Aleecia M McDonald, Robert W Reeder, Patrick Gage Kelley, and Lorrie Faith Cranor. A comparative study of online privacy policies and formats. In *International Symposium on Privacy Enhancing Technologies Symposium*, pages 37–55. Springer, 2009.

[125] John W McHoskey, Brian Hicks, Terri Betris, Chris Szyarto, William Worzel, Kristen Kelly, Tamara Eggert, Adam Tesler, Jenny Miley, and Travis Suggs. Machiavellianism, adjustment, and ethics. *Psychological Reports*, 85(1):138–142, 1999.

[126] McKinsey. How covid-19 has pushed companies over the technology tipping point—and transformed business forever. `https://mck.co/3trP4OV`, 2020.

[127] Patrizio Migliarini, Gian Luca Scoccia, Marco Autili, and Paola Inverardi. On the elicitation of privacy and ethics preferences of mobile users. In *Proceedings of the IEEE/ACM 7th International Conference on Mobile Software Engineering and Systems*, MOBILESoft '20, page 132–136, New York, NY, USA, 2020. Association for Computing Machinery.

[128] 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.

[129] Shervin Minaee, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu, and Jianfeng Gao. Deep learning–based text classification: A comprehensive review. *ACM Comput. Surv.*, 54(3), April 2021.

[130] Celia Moore, James R Detert, Linda Klebe Treviño, Vicki L Baker, and David M Mayer. Why employees do bad things: Moral disengagement and unethical organizational behavior. *Personnel psychology*, 65(1):1–48, 2012.

[131] Jill Mosteller and Amit Poddar. To share and protect: Using regulatory focus theory to examine the privacy paradox of consumers' social media engagement and online privacy protection behaviors. *Journal of Interactive Marketing*, 39:27–38, 2017.

[132] Aysen Gurcan Namlu and Hatice Ferhan Odabasi. Unethical computer using behavior scale: A study of reliability and validity on turkish university students. *Computers & Education*, 48(2):205–215, 2007.

[133] P. T. Nguyen, J. Di Rocco, D. Di Ruscio, A. Pierantonio, and L. Iovino. Automated classification of metamodel repositories: A machine learning approach. In *2019 ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems (MODELS)*, pages 272–282, Sep. 2019.

[134] Artur Nilsson. Humanistic and normativistic worldviews: Distinct and hierarchically structured. *Personality and Individual Differences*, 64:135–140, 2014.

[135] Artur Nilsson and John T Jost. Rediscovering tomkins' polarity theory: Humanism, normativism, and the psychological basis of left-right ideological conflict in the us and sweden. *PloS one*, 15(7):e0236627, 2020.

[136] Artur Nilsson and Samantha Sinclair. Death, ideology, and worldview: Evidence of death anxiety but not mortality salience effects on political ideology and worldview. 2021.

[137] Artur Nilsson and Michael Strupp-Levitsky. Humanistic and normativistic metaphysics, epistemology, and conative orientation: Two fundamental systems of meaning. *Personality and Individual Differences*, 100:85–94, 2016.

[138] Helen Nissenbaum. Privacy in context. In *Privacy in Context*. Stanford University Press, 2009.

[139] Tommaso Di Noia and Vito Claudio Ostuni. Recommender systems and linked open data. In Wolfgang Faber and Adrian Paschke, editors, *Reasoning Web. Web Logic Rules - 11th International Summer School 2015, Berlin, Germany, July 31 - August 4, 2015, Tutorial Lectures*, volume 9203 of *Lecture Notes in Computer Science*, pages 88–113. Springer, 2015.

[140] Patricia A Norberg, Daniel R Horne, and David A Horne. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs*, 41(1):100–126, 2007.

[141] Bernestein I Nunnally J. *Psychometric Theory.* McGraw-Hill, New York, 1994.

[142] 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. *Information, Communication & Society*, 23(1):128–147, 2020.

[143] Ernest H O'Boyle, Donelson R Forsyth, George C Banks, Paul A Story, and Charles D White. A meta-analytic test of redundancy and relative importance of the dark triad and five-factor model of personality. *Journal of personality*, 83(6):644–664, 2015.

[144] OECD. Oecd digital economy outlook 2020. `https://www.oecd-ilibrary.org/content/publication/bb167041-en`, 2020.

[145] Babatunde Tunde Ogunfowora, Viet Quan Nguyen, Piers Steel, and Christine C Hwang. A meta-analytic investigation of the antecedents, theoretical correlates, and consequences of moral disengagement at work. *Journal of Applied Psychology*, 2021.

[146] European Group on Ethics in Science, New Technologies, et al. *Statement on artificial intelligence, robotics and 'autonomous' systems: Brussels, 9 March 2018.* EU: European Union, 2018.

[147] Ernest H O'Boyle and Donelson R Forsyth. Individual differences in ethics positions: The epq-5. *PloS one*, 16(6):e0251989, 2021.

[148] Hae-Sang Park and Chi-Hyuck Jun. A simple and fast algorithm for k-medoids clustering. *Expert Systems with Applications*, 36(2, Part 2):3336–3341, 2009.

[149] The European Parliament and the Council of the European Union. Eu general data protection regulation (gdpr) - regulation eu 2016/679 of the european parliament and of the council of 27 april 2016. *Official Journal of the European Union*, 2016.

[150] Delroy L Paulhus and Kevin M Williams. The dark triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of research in personality*, 36(6):556–563, 2002.

[151] Lisa M Penney and Paul E Spector. Narcissism and counterproductive work behavior: Do bigger egos mean bigger problems? *International Journal of selection and Assessment*, 10(1-2):126–134, 2002.

[152] Marco Perugini and Luigi Leone. Implicit self-concept and moral action. *Journal of Research in Personality*, 43(5):747–754, 2009.

[153] Kostantinos V Petrides, Philip A Vernon, Julie Aitken Schermer, and Livia Veselka. Trait emotional intelligence and the dark triad traits of personality. *Twin Research and Human Genetics*, 14(1):35–41, 2011.

[154] Abinash Pujahari and Dilip Singh Sisodia. Modeling side information in preference relation based restricted boltzmann machine for recommender systems. *Information Sciences*, 490:126–145, 2019.

[155] A Pureesmali, M Molaee, Goradel J Alizadeh, and J Hashemi. *Dark triad personality and online self-disclosure in students.* Journal of psychological achievements (Journal of education & psychology), 2017.

[156] Mian Qin, Scott Buffett, and Michael W. Fleming. Predicting user preferences via similarity-based clustering. In Sabine Bergler, editor, *Advances in Artificial Intelligence*, pages 222–233, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[157] 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.

[158] EXOSOUL research team. Exosoul, the software exoskeleton. `https://exosoul.disim.univaq.it/`, 2021.

[159] Arnold Roosendaal. Digital personae and profiles as representations of individuals. In *IFIP PrimeLife International Summer School on Privacy and Identity Management for Life*, pages 226–236. Springer, 2009.

[160] Jenny Rosenberg and Nichole Egbert. Online impression management: Personality traits and concerns for secondary goals as predictors of self-presentation tactics on facebook. *Journal of computer-mediated communication*, 17(1):1–18, 2011.

[161] Norman Sadeh, Bin Liu, Anupam Das, Martin Degeling, and Florian Schaub. Personalized privacy assistant, March 23 2021. US Patent 10,956,586.

[162] Odnan Ref Sanchez, Ilaria Torre, Yangyang He, and Bart P. Knijnenburg. A recommendation approach for user privacy preferences in the fitness domain. *User Modeling and User-Adapted Interaction*, 30(3):513–565, July 2020.

[163] Elżbieta Sanecka et al. The dark side of social media: Associations between the dark triad of personality, self-disclosure online and selfie-related behaviours. *The Journal of Education, Culture, and Society*, 8(2):71–88, 2017.

[164] J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen. The adaptive web. pages 291–324. Springer-Verlag, Berlin, Heidelberg, 2007.

[165] Cynthia E Schairer, Cynthia Cheung, Caryn Kseniya Rubanovich, Mildred Cho, Lorrie Faith Cranor, and Cinnamon S Bloss. Disposition toward privacy and information disclosure in the context of emerging health technologies. *Journal of the American Medical Informatics Association*, 26(7):610–619, 2019.

[166] Cynthia E Schairer, Caryn Kseniya Rubanovich, and Cinnamon S Bloss. How could commercial terms of use and privacy policies undermine informed consent in the age of mobile health? *AMA journal of ethics*, 20(9):864–872, 2018.

[167] Adriano Schimmenti, Peter K Jonason, Alessia Passanisi, Luana La Marca, Nunzia Di Dio, and Alessia M Gervasi. Exploring the dark side of personality: Emotional awareness, empathy, and the dark triad traits in an italian sample. *Current Psychology*, 38(1):100–109, 2019.

[168] Shalom H Schwartz. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in experimental social psychology*, volume 25, pages 1–65. Elsevier, 1992.

[169] Kim Bartel Sheehan. Toward a typology of internet users and online privacy concerns. *The information society*, 18(1):21–32, 2002.

[170] Herbert A Simon. Models of man; social and rational. 1957.

[171] Glenn Simpson. Consumer-privacy issue turns a retired professor into a hot item. *WALL ST. J., June*, 25, 2001.

[172] Joginder P Singh. Managerial culture and work-related values in india. *Organization studies*, 11(1):075–101, 1990.

[173] Daniel J Solove. A taxonomy of privacy. *University of Pennsylvania law review*, pages 477–564, 2006.

[174] Eliane Sommerfeld. The subjective experience of generosity. 2010.

[175] Keith S Taber. The use of cronbach's alpha when developing and reporting research instruments in science education. *Research in science education*, 48(6):1273–1296, 2018.

[176] Mariarosaria Taddeo and Luciano Floridi. The debate on the moral responsibilities of online service providers. *Science and Engineering Ethics*, 22(6):1575–1603, 2016.

[177] Mariarosaria Taddeo and Luciano Floridi. How ai can be a force for good. *Science*, 361(6404):751–752, 2018.

[178] Steven Taylor, Ingrid C Fedoroff, William J Koch, Dana S Thordarson, Gary Fecteau, and Richard M Nicki. Posttraumatic stress disorder arising after road traffic collisions: Patterns of response to cognitive–behavior therapy. *Journal of consulting and clinical psychology*, 69(3):541, 2001.

[179] Silvan Tomkins. Left and right: A basic dimension of ideology and personality. 1963.

[180] Ludwig Toresson, Maher Shaker, Sebastian Olars, and Lothar Fritsch. Pisa: a privacy impact self-assessment app using personas to relate app behavior to risks to smartphone users. In *International Conference on Human-Computer Interaction*, pages 613–621. Springer, 2020.

[181] Beverley Townsend, Colin Paterson, TT Arvind, Gabriel Nemirovsky, Radu Calinescu, Ana Cavalcanti, Ibrahim Habli, and Alan Thomas. From pluralistic normative principles to autonomous-agent rules. *Minds and Machines*, pages 1–33, 2022.

[182] Onuralp Ulusoy and Pinar Yolum. Panola: A personal assistant for supporting users in preserving privacy. *ACM Transactions on Internet Technology (TOIT)*, 22(1):1–32, 2021.

[183] Jennifer M Urban and Chris Jay Hoofnagle. The privacy pragmatic as privacy vulnerable. In *Symposium on Usable Privacy and Security (SOUPS 2014) Workshop on Privacy Personas and Segmentation (PPS)*, 2014.

[184] David A Waldman, Danni Wang, Sean T Hannah, and Pierre A Balthazard. A neurological and ideological perspective of ethical leadership. *Academy of Management Journal*, 60(4):1285–1306, 2017.

[185] Yang Wang, Pedro Giovanni Leon, Kevin Scott, Xiaoxuan Chen, Alessandro Acquisti, and Lorrie Faith Cranor. Privacy nudges for social media: an exploratory facebook study. In *Proceedings of the 22nd international conference on world wide web*, pages 763–770, 2013.

[186] Jason Watson, Heather Richter Lipford, and Andrew Besmer. Mapping user preference to privacy default settings. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 22(6):1–20, 2015.

[187] Alan F. Westin. Privacy on and off the internet: What consumers want. Technical report, http://www.ijsselsteijn.nl/slides/Harris.pdf, 2002.

[188] Alan F. Westin. Bibliography of surveys of the u.s. public, 1970-2003, 2003.

[189] Alan F. Westin and Harris Louis Associates. Equifax-harris consumer privacy survey. tech. rep., 1996. conducted for equifax inc. 1,005 adults of the u.s. public. Technical report.

[190] Alan F. Westin and Harris Louis Associates. Harris-equifax consumer privacy survey. tech. rep., 1991. conducted for equifax inc. 1,255 adults of the u.s. public. Technical report.

[191] Jakob Wirth. Strength of ties as an antecedent of privacy concerns: a qualitative research study. In *Proceedings: Twenty-third Americas Conference on Information Systems, Boston, 2017*. AIS Electronic Library (AISeL), 2017.

[192] Pamela Wisniewski, AKM Najmul Islam, Bart P Knijnenburg, and Sameer Patil. Give social network users the privacy they want. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 1427–1441, 2015.

[193] Pamela Wisniewski, Heng Xu, Heather Lipford, and Emmanuel Bello-Ogunu. F acebook apps and tagging: The trade-off between personal privacy and engaging with friends. *Journal of the Association for Information Science and Technology*, 66(9):1883–1896, 2015.

[194] Pamela J Wisniewski, Bart P Knijnenburg, and Heather Richter Lipford. Making privacy personal: Profiling social network users to inform privacy education and nudging. *International Journal of human-computer studies*, 98:95–108, 2017.

[195] Ben Wolford. What is gdpr, the eu's new data protection law? `https://gdpr.eu/what-is-gdpr/`, 2020.

[196] H Woo, E Kang, Semyung Wang, and Kwan H Lee. A new segmentation method for point cloud data. *International Journal of Machine Tools and Manufacture*, 42(2):167–178, 2002.

[197] Allison Woodruff, Vasyl Pihur, Sunny Consolvo, Laura Brandimarte, and Alessandro Acquisti. Would a privacy fundamentalist sell their dna for $1000... if nothing bad happened as a result? the westin categories, behavioral intentions, and consequences. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*, pages 1–18, 2014.

[198] Virgil Zeigler-Hill, Amy E Noser, Courtney Roof, Jennifer Vonk, and David K Marcus. Spitefulness and moral values. *Personality and Individual Differences*, 77:86–90, 2015.

[199] Guoqiang Zhang, B. Eddy Patuwo, and Michael Y. Hu. Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting*, 14(1):35–62, 1998.

[200] Yanqi Zhao, Yong Yu, Yannan Li, Gang Han, and Xiaojiang Du. Machine learning based privacy-preserving fair data trading in big data market. *Information Sciences*, 478:449–460, 2019.

[201] Xingquan Zhu and Xindong Wu. Class noise vs. attribute noise: A quantitative study of their impacts. *Artif. Intell. Rev.*, 22(3):177–210, November 2004.

[202] Andrej Zwitter. Big data ethics. *Big data & society*, 1(2):2053951714559253, 2014.