Bridging the Gap Between GPDR and Software Development: The MATERIALIST Framework

Marco Saltarella
https://orcid.org/0000-0002-0021-9972
Giuseppe Desolda
giuseppe.desolda@uniba.it
https://orcid.org/0000-0001-9894-2116
Andrea Esposito
https://orcid.org/0000-0002-9536-3087
Francesco Greco
https://orcid.org/0000-0003-2730-7697
Rosa Lanzilotti
https://orcid.org/0000-0002-2039-8162

Research Article

Keywords: Privacy design patterns, GDPR, ISO 9241-210, code vulnerabilities

Posted Date: September 3rd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4273799/v2

Additional Declarations: The authors declare no competing interests.

Version of Record: A version of this preprint was published at Multimedia Tools and Applications on October 10th, 2024. See the published version at https://doi.org/10.1007/s11042-024-19923-0.

Bridging the Gap Between GPDR and Software Development: The MATERIALIST Framework

Marco Saltarella[1], Giuseppe Desolda[2*], Andrea Esposito[2], Francesco Greco[2], Rosa Lanzilotti[2]
marco.saltarella@finconsgroup.com
{giuseppe.desolda, andrea.esposito, francesco.greco, rosa.lanzilotti}@uniba.it
[1]*FINCONS SpA, Via Orfeo Mazzitelli, 258/E, 70124 Bari (BA), Italy*

[2]*Computer Science Department, University of Bari Aldo Moro, Via E. Orabona 4, 70125 Bari, Italy*

Abstract
As software production evolves, privacy is becoming an increasingly important consideration. This is especially true as national and supranational regulations, such as GDPR, require privacy as a mandatory aspect of software development. However, challenges such as a lack of knowledge about privacy and data protection regulations hinder the adoption of effective and compliant privacy implementation mechanisms. To address this issue, this article presents MATERIALIST, a methodological and technological framework that supports stakeholders involved in a software development lifecycle in including GDPR in their activities. Specifically, it provides design patterns that can be selected starting from GDPR articles, code vulnerabilities, and software lifecycle phases. The framework aims to facilitate the adoption of appropriate privacy implementation mechanisms in the software development lifecycle, thereby improving software quality.

Keywords
Privacy design patterns; GDPR; ISO 9241-210; code vulnerabilities.

1. Introduction

Cybersecurity threats are becoming increasingly dangerous, fuelled by the ever-connected nature of our society. Verizon's 2022 Data Breach Investigations Report found that 82% of breaches involve human elements [52], including social attacks, errors, and misuse. IBM's annual report also found that a data breach costs organizations approximately $4.35 million, and 83% of organizations are focusing on *when* a breach will happen rather than *if* it will happen [27].

Implementing security measures is thus crucial for protecting digital lives and upholding privacy regulations. In this direction, a comprehensive set of regulations for the gathering, handling, and disposing of personal data is outlined in the General Data Protection Regulation (GDPR), a landmark piece of legislation in the constantly changing field of data protection and privacy. The GDPR, which was enacted by the European Union in May 2018, has transformed how businesses handle personal data and strengthened people's rights to privacy. It has since become the industry standard for data protection.

Although the GDPR is supposed to be a sort of North Star guiding the development of secure software, designers and developers assigned to carry out its directives encounter noteworthy obstacles [42]. For example, it isn't easy to translate wide and frequently legally intricate GDPR articles into specific software development requirements [4]. Practices and solutions currently in use have had difficulty closing this gap. Ad-hoc and fragmented approaches have become commonplace in many organizations, leading to ambiguous and inconsistent software development and endangering data protection and privacy compliance [5, 26]. Moreover, according to Sobolewski et al., the GDPR is considered a step towards a more user-centric approach [47]. Still, due to the increased technological difficulties brought by the regulations, the usability of security and privacy mechanisms implemented for achieving compliance represents an open issue [33].

These shortcomings in how things are done now appear in several ways. One of the main problems with existing solutions is the lack of consistency in translating GDPR into software requirements [21]. Different interpretations and applications of the GDPR legal text can lead to inconsistent implementations, making it difficult for regulators to assess compliance accurately. Also, many current solutions do not comprehensively cover all articles and provisions of the GDPR, and this can leave critical aspects of data protection unaddressed, increasing the risk of non-compliance [26]. Moreover, translating regulatory language into technical requirements is a significant challenge: ambiguities, inconsistencies, and varying interpretations make it difficult for developers to extract precise, actionable guidance [4]. These problems are amplified by the fact that the interpretation of the GDPR articles might change over time, for example, in the light of previous judgments, new studies, but also the evolution of the technologies themselves (see the issue with ChatGPT, which was initially obscured in Italy because the guarantor deemed it non-compliant [31]. Existing solutions often struggle to keep up with these changes, leaving software development efforts vulnerable to obsolescence and regulatory misalignment [3, 15, 32].

In response to the shortcomings within the current body of literature and as a direct response to the multifaceted challenges outlined, this study aims to answer the following research question "How can software development be supported to ensure GDPR compliance?". To this end, this article contributes to the state-of-the-art by proposing MATERIALIST, a comprehensive methodological and technological framework that supports the integration of privacy aspects and GDPR in the software development lifecycle. Specifically, this framework empowers the selection and implementation of a set of 72 general-purpose Privacy Design Patterns (PDPs) and 15 user interface PDPs, which provide practical indications to design interfaces for GDPR privacy-related features. The PDPs selection can be performed starting from three distinct entry points, namely GDPR articles, code vulnerabilities, and phases of a human-centered software lifecycle process (ISO 9241-210 in this research), thus offering a practical, adaptable solution for developers to seamlessly integrate into their workflow.

MATERIALIST has also been successfully implemented as a web-based platform, meticulously crafted through a human-centered design process. Company designers, developers, and engineers have been

involved since the initial design of the platform to understand how stakeholders would seamlessly integrate this framework into their day-to-day practices. Moreover, a usability study was conducted involving key stakeholders to evaluate the usability and utility of this platform. The feedback and outcomes of this study played a pivotal role in refining and enhancing the platform, ensuring that it aligns more effectively with the real-world needs and expectations of the professionals who will rely on it.

The remainder of this article is organised as follows. Section 2 reports on the basics of the GDPR and the state of the art of privacy by design / privacy by default, and approaches to GDPR compliance. Section 3 describes the process that led to defining a set of new 15 Privacy Design Patterns for GDPR-Compliant User Interfaces. Section 4 reports the definition of the MATERIALIST framework, detailing the mapping phases between PDPs and GPDR articles, ISO 9241-210 phases and software vulnerabilities. Section 5 illustrates the design and implementation of the MATERIALIST web platform. Section 6 reports the details of the user study conducted to evaluate the usability e and utility of the platform. Finally, Section 7 draws conclusions and defines future work.

2. State of the art

2.1. Privacy and General Data Protection Regulation

Privacy is a fundamental right under the EU Charter of Fundamental Rights, defined as "the right of an individual to control the accuracy, use, and distribution of digital information about themselves" [49]. Privacy and security are intertwined, with security ensuring Authentication, Confidentiality, Integrity, and Availability of data [28]. Privacy results from controls over personal data processing [29], and it establishes how security mechanisms should be implemented, suitable conditions, and usage.

However, privacy is a complex aspect intertwined with social properties, involving personal autonomy, independence, power, control, and trust. It also raises moral, political, and legal questions [6]. Contrary to how careless online consumers are generally depicted about privacy, a study by Acquisti et al. [2] showed that users fundamentally care about their online privacy. Each person's sense of privacy is unique, and experiences and expectations may vary across domains of computing and society [1]. As a result, numerous laws and regulations have been enacted to regulate this field, making the issue of privacy a technical, social, and political concern that is difficult to untangle.

The General Data Protection Regulation (GDPR), published in the Official Journal of the European Union on May 4, 2016, and in force as of May 25, 2018, introduces stringent obligations and significant privacy challenges. It gives individuals ownership and control over their personal data, while organizations regulate data handling and introduce accountability at all stages of data processing. The European Commission aims to strengthen and harmonize personal data protection by simplifying the regulatory environment related to international affairs and unifying and standardizing privacy legislation within the EU. In essence, the GDPR protects user data in almost every way possible; it operates with the knowledge that data are the central engine upon which most businesses rely, and for this reason, it strives to protect that data at every stage of the process, giving users ultimate control over what happens to them according to the following principles, stated in Article 5:

- *Legitimacy*, fairness, and transparency: personal data must be "processed lawfully, fairly and transparently", which implies that the controller must have legitimate reasons for collecting the information and it must be transparent about which data are collected and for what purpose;
- *Purpose limitation*: personal data must be "collected for specific, explicit, and legitimate purposes" and used exclusively for the indicated purposes;
- *Data reduction*: personal data collected must be "adequate, relevant and limited to what is necessary";
- *Precision*: personal data gathered must be "precise and, if necessary, updated". The data subject has the right to request that the inaccurate data held are corrected or deleted;
- *Limitation on archiving*: personal data must be "stored in a form that allows identification of data subjects for no longer than necessary" and for the stated purposes;

- *Integrity and Confidentiality*: personal data must be "processed in such a way as to guarantee adequate security".

The GDPR is thus an EU-wide framework for safeguarding personal data, including digital and non-digital sectors. To comply, companies must securely manage consumer data and offer users access, control, update, and delete options. The GDPR also provides specific rights to data subjects, including the right to rectify inaccurate personal data (Art. 16), erasure their data without undue delay if there are no other impediments (Art. 17), request the restriction of processing when specific hypotheses are met (Art. 18), guarantee data portability (Art. 20), receive their data in a structured, commonly used, and machine-readable format (Art. 21), not to be subject to decisions based solely on automated processing, including profiling, that produces legal effects or significantly affects them (Art. 22).

Countries outside of the EU do not necessarily have a unified privacy regulation to comply with. For example, the United States does not have a comprehensive federal privacy law across the individual states. Moreover, regulations in force are sector-specific and vary according to the state laws, such as the California Consumer Privacy Act (CCPA) and the Health Insurance Portability and Accountability Act (HIPAA). Discrepancies in regulations reflect into differences in enforcement mechanisms and penalties. For instance, a non-compliant company may be subject to a fine of up to 4% of its global annual revenue, while in the U.S. penalties vary based on the specific state. Moreover, privacy law enforcement in the U.S. is often reactive rather than proactive, which can render the laws ineffective until after a data breach happens [43].

Moreover, organizations are mandated to accurately monitor user data processing activities and provide detailed information to authorities, enabling them to easily monitor and sanction organizations engaging in unlawful acts (Art. 30). The GDPR also mandates companies to protect data during processing, promoting techniques like pseudonymization, anonymization, and encryption. Companies must also ensure users control their data by incorporating safeguards to protect their rights. This compliance is crucial for businesses to maintain user privacy. The GDPR mandates companies to clearly inform users about data processing and obtain explicit and affirmative consent before proceeding with processing activities. Consent is defined as a freely given, specific, informed, and unambiguous indication of the data subject's wishes, indicating agreement to the processing of personal data. Ad-hoc requirements for consent conditions are introduced, with the data controller burdened with proof. Consent must be expressed in plain, comprehensible, and distinguishable language. The new GDPR rules require both the data controller and processor to be jointly responsible for data compliance, ensuring that both parties are liable for any violations resulting from outsourcing data entry or analysis to a third party. In addition, data-access policies should be regularly updated to address the potential for data misuse after consent is given. This is because techniques such as statistical inference and re-identification can be employed to extract information that the subjects did not initially consent to share [44].

The GDPR introduced a new professional figure called a Data Protection Officer who is mandatory for processing personal data by public authorities or public bodies, especially when operations require regular monitoring of data subjects on a large scale. The GDPR applies to organizations worldwide that offer services or products to European citizens, including those without a physical headquarters. If an organization is not based in the EU, it must appoint a company representative within the EU. Additionally, organizations must notify authorities in case of personal data breaches. GDPR conformance has become a global benchmark, with penalties including fines of up to 20 million euros or 4% of the organization's global turnover.

The introduction of GDPR raises new difficulties for business enterprises (both small and large) on four different levels: technical, legal, organisational, and regulatory [46]. On the technical front, GDPR requires the adoption of new techniques to process user data to ensure their privacy and to ensure their "right to be forgotten". Legally, the GDPR landscape is open-ended, making interpretation difficult. On the organisational level, GDPR mandates new roles that may not always be available (e.g., the Data Protection Officer). Regulatory-wise, the complexity of GDPR and the lack of basic support make it difficult to fully grasp its requirements [46]. The MATERIALIST framework, presented in this article, aims at alleviating these challenges, providing an interactive guide to understanding and navigating GDPR for a successful implementation.

2.2. Privacy by design and privacy by default

The GDPR, as stated in Article 25, mandates the implementation of specific data protection mechanisms through the use of privacy by default and privacy by design principles. *Privacy by default* aims to ensure privacy-preserving configurations, while *privacy by design* embeds privacy into systems from the beginning. Specifically, Article 25 states:

"1. Taking into account the state of the art, […], the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organizational measures, [...] which are designed to implement data-protection principles, […], in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects. 2. The controller shall implement appropriate technical and organizational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed."

Ann Cavoukian pioneered the concept of privacy by design, defining the seven core principles of privacy by design in one of her most famous research [16]:

1. *Proactive not Reactive – Preventative not Remedial*, which asks for striving to avoid any potential privacy issues and being proactive in this regard rather than just waiting for the risk to materialize;
2. *Privacy as the Default Setting*, which implies that the system should be set up by default with the privacy-preserving configuration;
3. *Privacy Embedded into Design*, which suggests that privacy should be considered as a process across the whole system's design process and not simply at a single, arbitrary timeframe.
4. *Fully Functionality – Positive-Sum, not Zero-Sum*, that implies that engineers should make an effort to take into account all of the interests of the system's concerned stakeholders to optimize both functionality and privacy;
5. *End-to-End Security – Full Lifecycle Protection*, which implies that data should be secured during their whole lifespan, from ingestion to deletion;
6. *Visibility and Transparency – Keep it Open*, which asks for being transparent about personal data processing to enhance the system's trustworthiness;
7. *Respect for User Privacy – Keep it User-Centric*, which suggests protecting the involved users' privacy as one of the ultimate objectives. Thus, a user-centric strategy should be used to assist users in making an informed decision about their privacy.

It is noticeable how the GDPR draws on these principles in one form or another: for instance, Article 5 of the GDPR requests for the integrity and confidentiality of data (Full Lifecycle Protection) as well as lawfulness, fairness, and transparency (Keep it Open). These guidelines constitute the foundation of the Privacy by Design methodology, which seeks to incorporate privacy and security into ICT (Information and Communications Technology) operations and architectural designs. According to this paradigm, privacy should be regarded as a fundamental component of information systems, demanding the incorporation of security and privacy into the system design process.

Privacy by Design is inextricably linked to the concept of Security by Design, an older paradigm based on the CIA (Confidentiality, Integrity, and Availability) principles, which state that only authorized parties should be allowed to access information (confidentiality), data should be protected against unauthorized change or erasure (integrity), and data should always be available when requested by an authorized party (availability).

Cavoukian merges privacy and security by default into a single concept [17], promoting a re-evaluation of system design and integrating privacy by default as technology, user expectations, and regulations evolve [18]. The privacy-by-default and privacy-by-design paradigms are closely linked, requiring the implementation of security policies to ensure privacy-preserving configurations. These policies follow the "least" principles, minimizing authorization privileges, trust, and information access to mitigate security risks. Cavoukian's work has been a significant turning point in literature, providing a solid foundation for our current regulatory system. However, understanding how these approaches can be

applied in everyday practices is crucial, as they suggest only high-level requirements during system design processes.

2.3. Privacy compliance methodologies

As previously said, privacy regulations such as GDPR mandate that software be created with specific security and privacy safeguards to protect users and their data, as well as to ensure the secure administration of the entire personal data process.

Various frameworks and approaches in the literature address privacy and security issues in software development. P-STORE is a 10-step software development technique that begins with data privacy goals and ends with elicited privacy needs [48]. SQUARE, or Security Quality Requirements Engineering, is another concept. It is a 10-step method that considers the security concerns of all stakeholders engaged in the software system [37]. Microsoft also recommends its Secure Development Lifecycle (SDL), which states that creating a secure software system is critical to consider an attack model and how an attacker can compromise the system by exploiting design flaws and, as a result, to implement the necessary defense mechanisms in the system.

Brodin presents a framework for GDPR compliance in small and medium-sized businesses, although the framework does not focus on how the legislation would be technically implemented in software [13]. The H2020 project BPR4GDPR, which proposes an 8-phase framework for enabling the re-engineering of business processes to ensure compliance by design, takes a more technical approach [34], with a compliance toolkit that combines several functionalities, ranging from notification systems to cryptography to aid in the implementation of privacy-aware systems.

Riva et al. analyze existing methodologies to address GDPR and privacy-by-design principles, categorizing them into three categories - pattern-based, model-driven, and framework-oriented - stating that GDPR principles should be wholly integrated into the software development lifecycle to ensure compliance, although there is still no unified and comprehensive framework, and thus more work needs to be done to propose more practical and holistic methodologies to integrate privacy requirements into the software development lifecycle [41].

Baldassarre et al. present a privacy knowledge base and a related visual tool to assist developers during the development process, which are based on five key concepts of privacy and security engineering (privacy-by-design principles, privacy design strategies, privacy patterns, software vulnerabilities, and applicative context) [8].

The study described in [14] provides a thorough investigation of business process awareness concerning compliance with the General Data Protection Regulation (GDPR). It examines the intersection of business processes and GDPR requirements, highlighting the difficulties and prospects for organizations aiming to comply with GDPR regulations. The authors offer valuable insights into the importance of integrating business process awareness to effectively support GDPR compliance initiatives. The paper provides guidance for organizations navigating the intricate landscape of data protection regulations by addressing the complexities of GDPR compliance within the realm of business processes. It contributes to the wider discussion on GDPR compliance and its impact on business operations, making it a relevant reference for related studies in the areas of compliance, business processes, and GDPR adherence.

However, none of the works presented here provide a tool to support each approach's phases or place a heavy emphasis on usability difficulties. There is less emphasis on holistic approaches in this respect and more on contributions to individual areas, such as usable GDPR-compliant dashboards [22, 39], privacy policies [40], opt-out, and data deletion [23]. Indeed, as previously stated, privacy should be secured not only through technological methods, but because this is a multidisciplinary issue, Human-Computer Interaction (HCI) techniques should be considered at every stage of system implementation. While it is acknowledged that engineering security and privacy throughout the development process is critical, there is now also a need to ensure that the system is developed with the appropriate usability requirements to guarantee proper and secure usage by users [38], as users frequently have

misconceptions about the perceived privacy risks when using online services [7]. Furthermore, when designing a system, practitioners must keep in mind the aim of the users' interaction with the system, how privacy-preserving technologies are integrated, and the extent to which these security mechanisms can interfere with users' goals and potentially put the user in danger if ignored or bypassed [33].

In this regard, HCI can play an important role in assisting and designing systems that safeguard user privacy in the most frictionless manner possible [25]. Thus, as we will see in the following, this might be accomplished by including HCI methodologies across the software development lifecycle, as well as Privacy by Design and Default approaches, to bring to light a technical and practical privacy-understanding from the user's perspective [50]. However, according to Shneiderman et al., designing for usable privacy and security is still one of the key issues that HCI researchers confront to ensure users' digital lives are secured [45]; therefore, more work in this direction is undoubtedly required.

MATERIALIST integrates the HCI perspective by proposing a set of PDPs focused on user interface and user interaction. Moreover, it drives the selection of PDPs starting from all phases of the most important human-centered software lifecycle, namely ISO 9241-210. Of course, we do not intend to leave out more implementation aspects of the software that must comply with the GDPR; the MATERIALIST entry points are also mapped with a set of more technical and already existing 72 patterns (see the next section). We have also defined architectural and programming patterns, and we have mapped them with the MATERIALIST entry points, but this is beyond the scope of this manuscript, which focuses on usable interaction aspects of GDPR-compliant software.

3. Definition of a set of Privacy Patterns to Design GDPR-compliant User Interfaces

Design patterns are recurring solutions to common problems. In general, they are defined by domain experts and provide structured and reusable knowledge [20]. In software development, they assist developers with avoiding the "reinventing the wheel" error, i.e., not re-creating solutions that already exist from scratch, but instead using generally acknowledged and tested solutions. This makes the developers' job easier and faster since developers have to identify the best pattern and implement it by contextualizing it to their own needs. A common language and structure are employed in design patterns to reduce or, better yet, eliminate misconceptions and enable the practical execution of the offered solution.

We can find more specific patterns in different contexts. For example, with reference to the scope of our research, in cybersecurity, Privacy Design Patterns (PDPs) are often adopted to address common privacy issues; they can be viewed as a manner of turning the privacy-by-design concept into viable software engineering practice [19]. One of the most important examples is the catalog of PDPs resulting from a joint research work between the U.S. Department of Homeland Security and the National Institute of Standards and Technology. This catalog is available at https://privacypatterns.org/ and consists of 72 privacy patterns grouped by the Privacy Design Strategies (control, abstract, separate, hide, minimize, inform, enforce, demonstrate). Similarly, User Interface design patterns ease navigation and reduce cognitive load during the interaction of a user with a system [11]. However, although the privacy of software often involves aspects of the user interface (UI) and user interaction, to the best of our knowledge, no PDPs for user interface exist.

To fill this gap, one of the contributions of this research consists of defining a new set of PDPs for the UI. The definition of this new catalog of PDPs followed a 4-step protocol. It aimed to extend the catalog of the 72 PDPs available at privacypatterns.org with more specific patterns that address UI aspects. For this work, we selected a subset of 5 patterns from the 72, choosing from those we found most impactful and critical for the UI. Specifically, the selected patterns are:

- Appropriate Privacy Feedback
- Privacy Dashboard
- Layered Policy Design
- Privacy Labels

- Trust Evaluation of Services Sides

The remaining PDPs that involve UI aspects will be treated as an extension of this work. In the following, we detail the process with some explicative examples of how it has been followed.

3.1. Identification of keywords

For each PDP that acted as a seed for defining a specific PDP for UI, the starting point of the entire process is a literature review of the existing solutions related to the privacy aspects the pattern is dealing with. To ensure comprehensive coverage of the articles in literature, a set of query strings is defined starting from the following structure:

- privacy [topic] user interface
- privacy [topic] UI
- GDPR [topic] user interface
- GDPR [topic] UI

with [topic] word in the brackets replaced with the keyword(s) representing the pattern's main concept. For example, considering the possible integration of the pattern *Appropriate Privacy Feedback*[1] with UI aspects, the following queries are used:

- privacy feedback user interface
- privacy feedback UI
- GDPR feedback user interface
- GDPR privacy feedback UI

3.2. Execution of the research queries and articles' selection

After that, the search strings are typed into Google Scholar to discover relevant scientific publications. For each result, the title and abstract are read to understand if the article fits the research and note after how many results the articles begin to become irrelevant to our scope. For example, for the search string "privacy feedback user interface" Scholar returned 772,000 results, but only the top 12 results seemed relevant.

Then, we performed a final selection of the relevant articles based on 3 inclusion criteria:

- Peer reviewed
- Written in English
- Focussed entirely on the theme of the pattern (e.g., feedback)

3.3. Articles clustering

The articles selected in the previous phase may cover very different parts of the same topic, such as feedback on different devices (mobile/desktop), synchronous and asynchronous feedback, etc. In this phase, the articles are classified into sub-groups covering the same topics. Each subset of articles contributes to the definition of a new specialized UI pattern.

Three of the authors of this article thoroughly review and analyze all clusters to discover the strengths of the solutions offered in the articles to be reported in the pattern. The criticality lies, for example, in the fact that the articles contribute to formulating the pattern in very different ways, some with specific answers and others with higher-level solutions such as principles, lessons learned, and guidelines.

[1] https://privacypatterns.org/patterns/Appropriate-Privacy-Feedback

3.4. Patterns Definition

The patterns are written and presented according to the model provided by PDPs available at https://privacypatterns.org, which is among the ones most used in the literature. Figure 1 shows the structure of a PDP.

Title: Title of the pattern.

Summary: A short overview of the pattern

Context: A description of the general situation under which the problem occurs and where the pattern may be applied, including possible examples. It describes the scope, market, user, or other conditions that, if changed, would alter the problem or solution.

Problem: The problem that repeatedly occurs and the pattern addresses, including issues related to users, computing systems, legal regulation, etc.

Solution: The solution the pattern suggests solving the problem. It should allow the reader the freedom to craft the solution most appropriately. This section can also include the following two sub-sections:

- [*Structure*]: The configuration of the elements that solve the problem and how they interact
- [*Implementation*]: A description of the possible technical implementation
- *Consequences*: Benefits and potential disadvantages of the pattern, including reference to the GDPR articles addressed if relevant.
- [*Constraints*]: Any constraints to be considered for the application of the pattern

Examples: Explicative examples of the application of the pattern

- Known Uses: Link and/or screenshot to known uses of the pattern
- Related Patterns: List of other strictly related privacy patterns, that may support the implementation of this pattern

Sources: Sources of the articles must be reported and referenced in the description of the pattern.

Figure 1. Structure of a Privacy Design Pattern. The sections in square brackets are optional.

The execution of this protocol led to the definition of 15 UI PDPs divided into the following categories:

- *Feedback*: grouping patterns focused on privacy-related feedback to the user with respect to different aspects, such as access to sensitive data (e.g., geo-location), the timing of the feedback, and context (e.g., mobile, desktop). The newly defined patterns in this category are: Mobile Synchronous Feedback, Feedback in dashboards to manage privacy for mobile applications, Feedback in dashboards to manage desktop privacy, Timed Asynchronous Feedback on mobile, Feedback on geographic location, Social media feedback;
- *Passwords*: grouping patterns focused on UIs to support users in selecting a secure password, including new authentication mechanisms such as graphical and 3D passwords. The newly defined patterns in this category are: Textual passwords, Graphical passwords, 3D passwords;
- *Privacy Icons*: grouping patterns focused on providing guidelines on privacy-related icon design and usage. The newly defined patterns in this category are: Privacy Icons for desktop, Privacy Icons for mobile;
- *UI for access control*: grouping patterns focused on providing guidelines to design interfaces handling access control features to online resources. The newly defined patterns in this category are: UI for web access control, UI for mobile access control;
- *Privacy Labels*: grouping patterns focused on providing guidelines on privacy-related label design. The newly defined patterns in this category are: Privacy Labels for Apps' Stores, Privacy Labels for Web Applications.

In the remainder of this section, we report the pattern "Privacy Labels for Apps Stores" as an example. The complete definition of the rest of the patterns is available in the Knowledge Base section of the MATERIALIST platform (see Section 5).

Title: Privacy Labels for Applications' Stores

[Also Known As]

Privacy Nutrition Labels for Applications' Stores

Summary: Software developers' responsibility in dealing with the ever-growing privacy requirements from platform providers (e.g., Apple and Google) has increased since the enaction of prominent laws and regulations – i.e., GDPR and CCPA. The main goal of this pattern is to empower users to learn about the app's collection and use of data before installation (Li et al., 2022). Privacy visualizations – visual representations designed to communicate aspects related to the handling of personal data to users of online services (Barth et al., 2022) – come in handy in the shape of Privacy Nutrition Labels which, for what concerns Application Stores and Market Places, represent just an emerging, not fully explored, and recently debated topic.

With this pattern, we intend to provide design opportunities as well as a general framework for Privacy Nutrition Labels that are suitable for Applications' Stores and Market Places to help both users and developers understand – the firsts – and design – the seconds – these labels in an intelligible and correct manner.

Context: People daily download and use various applications that may affect their privacy. In general, each application, depending on the services they offer, and on developers' purposes, may differ in the collection and gathering of data. Thus, users must be aware of privacy details before downloading an application.

Developers, on the other hand, should provide a clear, uniform, and brief summary of what data is collected along with how it is used and shared (Li et al., 2022) to complement the mandatory but cumbersome privacy policies. Indeed, they are frequently not read because they are typically difficult and are often too complicated, long, and confusing.

Problem: Past research has shown that most users do not pay attention to the data that may be required by an app before its installation (Scoccia et al., 2022). Users need to be informed in an intelligible way about the application's privacy policies, directly on the Application Store so that they can evaluate the potential consequences of their consent and choices, and about the collection and use of their data before installation. Although it has been established, by numerous studies, that people do not read privacy policies, they often don't believe they have a choice when it comes to their privacy and the use of their personal information (Kelley, 2009). Moreover, reading current online privacy policies is challenging and time-consuming (Kelley et al., 2010); this is mainly due to the use of specific terms that lead people to confusion in understanding how requested data are related to the use of the app (Kelley, 2009).

Software developers have increasing responsibility to deal with the ever-growing privacy requirements from platform providers – e.g., Apple, Google, etc. – and recently enacted laws – i.e., GDPR – and consequently face an increasing number of challenges (Li et al., 2022).

Forces and Concerns:

- Users want to understand the risks in as short time as necessary, at a granularity most suitable to them;
- Users would rather be provided with relevant and ideally concise information rather than cumbersome privacy policies;
- Users want to be informed of the choice they have in managing their information, in particular: whether specific information requests are optional, if sharing can be limited, and if it is possible to request access and changes to their information or have it deleted (Kelley, 2009);

- Application Stores and Market Places want to assist developers in notifying consumers about what information will be collected, how it will be used, and with whom it will be shared before the installation of an application to help users make an informed decision (Kelley, 2009);
- Developers want to ensure that users are fully aware of the risks of using the system before installing and using it.

Solution: Provide users with knowledge of data collected by the application directly on the Application Store using simple privacy nutrition labels that display the most crucial aspects of the privacy policy, which users are most likely to read. Nest a successive level of detail so that users can quickly deepen their knowledge about information that is relevant to them if needed.

[*Structure*]: To improve the visualization of Privacy Nutrition Labels for Applications' Stores and give people a different level of detail (Johansen et al., 2022), a possible structure can have two main layers (Li et al., 2022):

- On the first layer, users can see the high-level data categories that the app collects (e.g., location) and whether this data category is linked to users or used to track users. Users can click the see details link to explore the second layer.
- The second layer includes more specific data types – e.g., Coarse/Precise location –, what the data is used for – e.g., Third-Party Advertising, App Functionality –, and whether each data type is linked to users or used to track users.

The goal of Layered Privacy Nutrition Labels is to give the user a bird's-eye view of which privacy aspects are included in the privacy policy. Then, if the user wants more information, they can drill down to a deeper level. At the bottom layer, one can find the link to the full privacy policy (Johansen et al., 2022).

In the following, we present some guidelines to implement this pattern:

- Use bold rules to separate sets of information, giving the reader an easy roadmap through the label and clearly designating sections that can be grouped by similarity (Kelley et al., 2009);
- Each block should provide a clear and boldfaced title (Kelley et al., 2009) – e.g., Data Linked to You or Data Used to Track You (Li et al., 2022) – followed by a short textual description (Kelley et al., 2009);
- Pair both the title and each label with standardized icons to communicate related privacy aspects. Icons may be grouped into different categories based on whether they represent how data are used and stored, in order to provide a visual summary of the most important privacy-related information (Barth et al., 2022). In order to ensure readability and ease of translation between platforms, icons could be implemented through standard emojis (McParlan & van der Linden, 2021).

Since policies are too long to read, the nutrition label must only present the most relevant information for users, e.g., data collection, data sharing, security, and period of retention [36].

[*Implementation*]: iOS developers can – and must – exploit a web-based tool, App Store Connect, provided by Apple, to be guided in the filling of relevant information that will be automatically inserted and nested in the Privacy Nutrition Labels. Consequently, the design of each label is automatically provided by Apple. All the privacy details are, indeed, self-reported by means of this tool.

As for what concerns other Application Stores, they're not currently provided by automated mechanisms, and they need to implement such a framework to be consistent with the new privacy regulations.

In order to implement such a pattern, we highly suggest following Apple's implementation, which can be considered a milestone of Privacy Nutrition Labels, while applying the recommendations specified in this pattern.

Consequences: This section contains an overview of the positive and negative impacts regarding the introduction and application of this pattern.

The positive impacts of Privacy Nutrition Labels are the following (Kelley et al., 2009; Li et al., 2022):

- Studies' results reveal that the nutrition label made it easier to compare between a small set of items, allowing consumers to benefit, through informed decision-making.
- Labels presented clearly and at relevant times could affect users' decisions when choosing between similar apps.
- Requiring developers to offer a more succinct, readable privacy notice may incentivize them to adopt privacy-friendly designs, since collecting less data will make creating privacy labels easier.
- Privacy labels may lead to more awareness and thus to a voluntary reduction in data sharing.

The negative impacts of privacy labels are the following:

- If many labels are inaccurate, it may lead to mistrust by users and hinder long-term adoption (Li et al., 2022).

Example: User X browses the Application Store searching for an app that fits their needs. User X finds an application that seems suitable, but they are concerned about what data the app will collect about them. Before installing it, they scroll down on the app page (Figure 2-left) and find a section dedicated to the app's privacy practices, the Application Privacy Section, in which one or more privacy nutrition labels are displayed (Figure 2-center). X can click on a nutrition label, and a new window appears showing them the app's privacy policies in more detail (Figure 2-right).

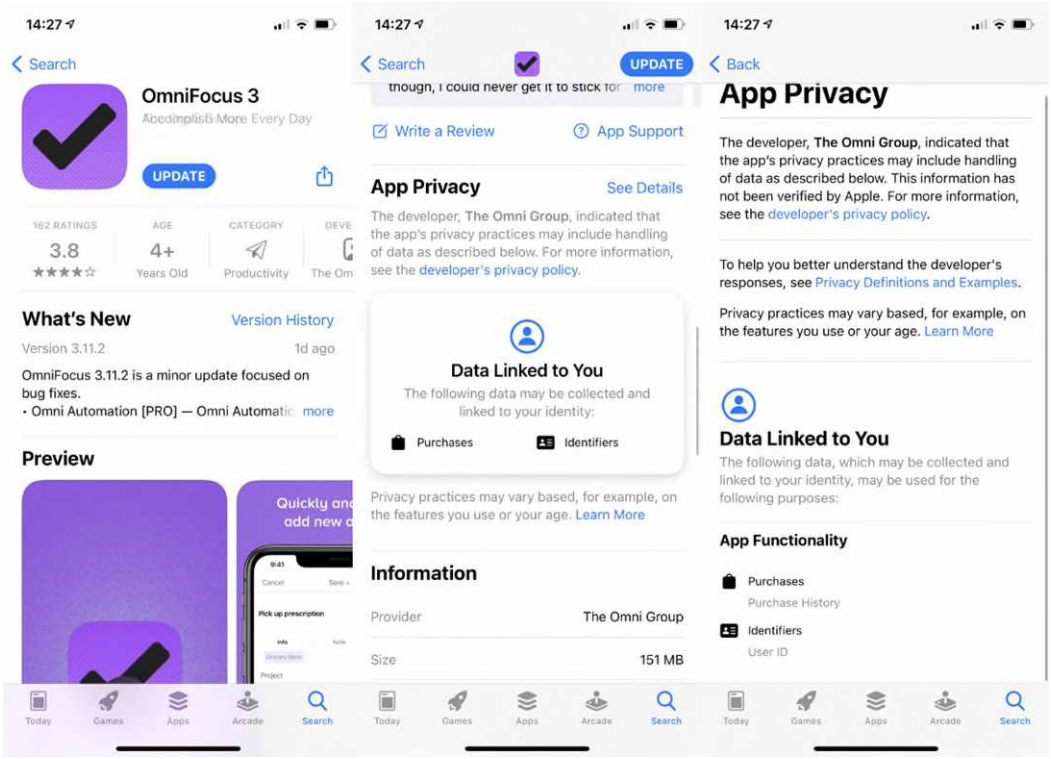
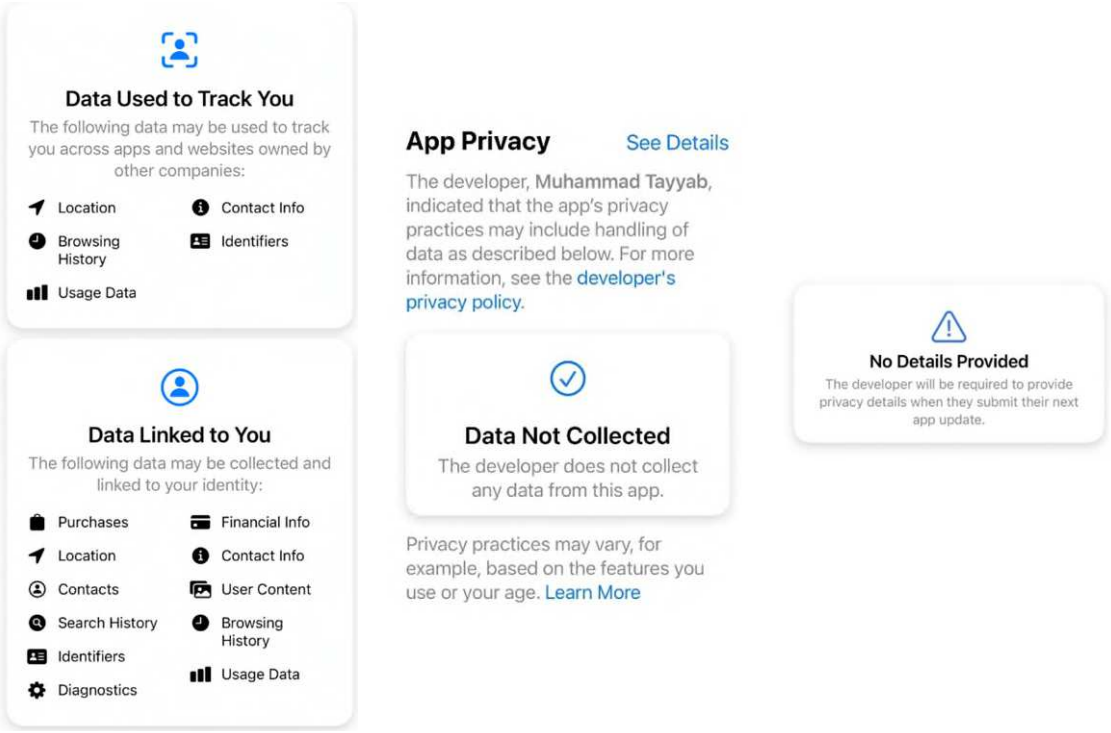


Figure 2. Use case example of Privacy Nutrition Labels for Applications Stores: app page (left), first layers (center), second layer (right).

[*Known Uses*]: As of December 2020, Apple requires all apps to provide app privacy details, which the Apple app store displays as a privacy label in an App Privacy section on each app's product page to empower users to learn about the app's collection and use of data before installation.

To provide the app's privacy details, Apple decided to adopt the Privacy Nutrition Labels, providing multiple standard labels like the following:

- Data collected Used to Track the User (Figure 3-a top).
- Data collected Linked to the User (Figure 3-b bottom).
- No Data collected (Figure 3-c).



a b c

Figure 3. Apple's Privacy Nutrition Labels when data are collected (a); Apple's Privacy Nutrition Labels when data are not collected (b); Apple's Privacy Nutrition Label when Developer decides to do not provide privacy information (c).

Sources:

- Barth, S., Ionita, D., & Hartel, P. (2022). Understanding Online Privacy—A Systematic Review of Privacy Visualizations and Privacy by Design Guidelines. ACM Comput. Surv., 55(3). https://doi.org/10.1145/3502288

- Johansen, J., Pedersen, T., Fischer-Hübner, S., Johansen, C., Schneider, G., Roosendaal, A., Zwingelberg, H., Sivesind, A. J., & Noll, J. (2022). A multidisciplinary definition of privacy labels. *Information & Computer Security*, *30*(3), 452–469. https://doi.org/10.1108/ICS-06-2021-0080

- Kelley, P. G. (2009). Designing a privacy label: Assisting consumer understanding of online privacy practices. CHI '09 Extended Abstracts on Human Factors in Computing Systems, 3347–3352. https://doi.org/10.1145/1520340.1520484

- Kelley, P. G., Bresee, J., Cranor, L. F., & Reeder, R. W. (2009). A "nutrition label" for privacy. *Proceedings of the 5th Symposium on Usable Privacy and Security - SOUPS '09*, 1. https://doi.org/10.1145/1572532.1572538

- Kelley, P. G., Cesca, L., Bresee, J., & Cranor, L. F. (2010). Standardizing privacy notices: An online study of the nutrition label approach. *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*, 1573. https://doi.org/10.1145/1753326.1753561

- Li, T., Reiman, K., Agarwal, Y., Cranor, L. F., & Hong, J. I. (2022). Understanding Challenges for Developers to Create Accurate Privacy Nutrition Labels. CHI Conference on Human Factors in Computing Systems, 1–24. https://doi.org/10.1145/3491102.3502012

- McParlan, J., & van der Linden, D. (2021). Privacy labels should go to the dogs. Eight International Conference on Animal-Computer Interaction, 1–10. https://doi.org/10.1145/3493842.3493888

- Scoccia, G. L., Autili, M., Stilo, G., & Inverardi, P. (2022). An empirical study of privacy labels on the Apple iOS mobile app store. *Proceedings of the 9th IEEE/ACM International Conference on Mobile Software Engineering and Systems*, 114–124. https://doi.org/10.1145/3524613.3527813

4. The MATERIALIST Framework

The catalog of 15 PDP for UI, mentioned in the previous section, represents the first milestone of this research that aims to support developers and engineers in integrating privacy solutions in their user interfaces. However, this catalog alone is not enough because stakeholders might have difficulty in choosing the right PDP, especially if they have to start from specific GPDR articles. For this reason, some authors of this article already proposed a preliminary version of the framework called MATERIALIST [3], which mapped the 72 privacy design patterns found on privacypatterns.org and some of the UI patterns (at that time, only 5 PDP for UI have been proposed, the remaining ones are part of this work) with: i) GDPR articles, ii) ISO 9241-210 phases, and iii) privacy vulnerabilities found during static code analysis (the OWASP Top 10 2021, de-facto standard, is used to identify vulnerabilities). One unique aspect of this framework is the ability to traverse between entry points using privacy patterns, allowing stakeholders to easily identify PDPs to apply to different phases of development or specific vulnerabilities. In forward engineering, PDPs are suggested to comply with GDPR articles, while in backward engineering, patterns are indicated for each ISO 9241-210 phase to support stakeholders throughout the development process. In the following subsection, we detail the mapping of the MATERIALIST PDPs and the three entry points (see Figure 4).

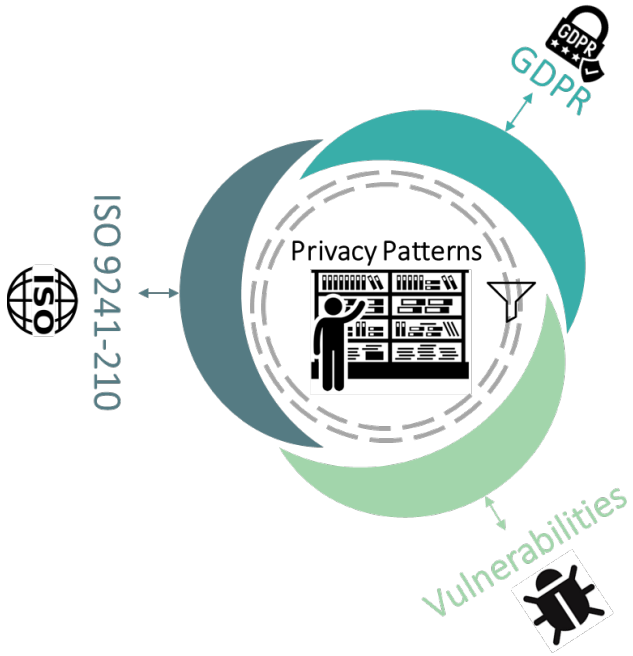


Figure 4. MATERIALIST mapping between PDPs and the three entry points

4.1. Mapping between GDPR and PDPs

As mentioned before, one of the most important concerns to consider when complying with data privacy regulations is that they only provide indications in legal language. To overcome this constraint, PDPs can support the implementation of privacy regulations and encourage privacy by design and default practices, as explicitly mandated by GDPR article 25. GDPR, on the other hand, does not provide any guidance on how to comply with the privacy rules, leaving developers with no clear references for the regulation. Furthermore, the GDPR has 99 articles, the bulk of which are merely legal and administrative, necessitating a focus on the articles that may have repercussions for the system's evolution.

We performed a comprehensive mapping between privacy patterns and GDPR articles to bridge this gap and assist stakeholders in selecting the proper privacy design patterns in accordance with the GDPR articles with which they must comply. Two researchers with experience in both PDPs and GDPR worked independently on this endeavor and then discussed their findings to reach an agreement on the mapping. To improve robustness and reduce the risk of biases, this mapping was performed starting from i) each PDP towards the GDPR articles, and ii) from each GDPR article towards the PDPs. The mapping technique, in particular, was first carried out independently by the two researchers in both directions. The researchers spent roughly 40 hours on this step. They then compared their findings and discovered an interrater reliability of 67%. Finally, they talked about the differences and came to an agreement on the remaining mappings.

This process discovered numerous relationships between each PDP and 14 different GDPR articles, indicating which PDPs may enable compliance with them. Due to the GDPR's lack of financial incentives for user data sharing, three patterns – "Incentivized Participation," "Pay Back," and "Reciprocity" - remain unassigned. The complete result of this mapping is reported in Table 1 in the appendix and it has been integrated into the MATERIALIST platform.

To this end, stakeholders can use this mapping as part of the broader MATERIALIST framework to help them identify the appropriate PDPs starting with the GDPR articles they need to be compliant with, and vice versa can understand what GDPR article compliance is supported when implementing a specific PDP.

4.2. Mapping between ISO 9241-210 and PDPs

The ISO 9241-210 standard recommends a complementary and iterative process for software development that is divided into four iterative phases that are cycled until the solution meets all of the user requirements:

- *Understanding and specifying the context of use*: the execution environment (both physical and/or digital), the tasks and goals of the end users, as well as their characteristics, must be explored to identify potential challenges, needs, and restrictions of the future system;
- *Specifying the user requirements*: express both functional and non-functional needs, and provide an explicit link between user requirements and the context identified in the preceding stage, as well as the system's business aim, resolving and documenting trade-off decisions between requirements;
- *Producing design solutions*: design tasks, interactions, and user interfaces based on defined needs and feedback from earlier iterations, using state-of-the-art, guidelines, standards, prototypes, simulations, and/or mock-ups;
- *Evaluating the design*: offer feedback and compare design ideas from the user's perspective to better understand users' demands and assess requirement achievement and standard conformity.

A mapping is also performed between each PDP and all ISO 9241-210 process phases. This mapping tries to identify the phase in the development process when stakeholders should start thinking about implementing PDPs. This is critical because security must be considered as a process that spans the software development lifecycle. Furthermore, despite the popular belief that patterns should only be considered during the coding process, their implementation is heavily dependent on the context and system definition. The pattern Data Breach Notification[2], for example, should be considered in the requirements phase because the GDPR, and similarly other regulations, demand that authorities and users be notified of the incident within 72 hours. Special requirements and business processes must thus be defined to properly apply this design and thereby comply with the legislation. A pattern, on the other hand, like the Privacy Aware Network Client[3], which focuses on architectural and technological challenges, belongs to the design production phase; it is thus easier to design a safe system thanks to meticulous planning, made possible by engineers knowing when to begin evaluating a pattern. Furthermore, because usability is sometimes overlooked in security and privacy concerns, implementing a human-centered development approach such as ISO 9241-210 [30] aids in the implementation of more usable solutions.

The complete mapping between the ISO 9241-210 stages and PDPs is reported in Table 2 in the appendix and it has been integrated into the knowledge base section of the MATERIALIST platform.

4.3. Mapping between Vulnerabilities and PDPs

Concerning the vulnerabilities, the mapping conducted refers to prior work by Baldassarre et al. [9] but is updated to take into account OWASP Top 10 2021 rather than OWASP Top 10 2017. It is important to highlight that there is no direct relationship between vulnerabilities and patterns; rather, vulnerabilities are first mapped with privacy-by-design principles, which are then mapped with PDPs via privacy design strategies. The mapping is described in [10] and is included in the MATERIALIST tool's knowledge base.

5. Fostering the adoption of MATERIALIST in real contexts: a web platform

A web platform that supports stakeholders using the MATERIALIST framework in real scenarios was developed. It is the result of a human-centered design process: company designers, developers, and

[2] https://privacypatterns.org/patterns/Data-breach-notification-pattern

[3] https://privacypatterns.org/patterns/Privacy-aware-network-client

engineers have been involved to collect important indications on how stakeholders would effortlessly integrate MATERIALIST into their daily jobs. This platform has also been evaluated during a usability study to evaluate the integration of the web platform into the software development lifecycle of a company. The study's findings were crucial in helping to improve and refine the web platform so that it better meets the demands and expectations of the professionals who will use it in the real world. The following sub-sections present the design and the implementation of the MATERIALIST web platform, while its evaluation is reported in Section 6.

5.1. Platform design: an elicitation study

An elicitation study was carried out to design a tool that users may use effectively to select patterns according to the MATERIALIST framework. Indeed, fulfilling the needs of the system's various stakeholders and obtaining a shared understanding and agreement on the given requirements is one of the most challenging tasks in software development [24]. To that purpose, a five-step methodology was followed in accordance with the instructions of [51] to collect the requirements of the MATERIALIST tool. In particular, a semi-structured interview model was chosen to perform the elicitation step. This process usually starts with some ice-breaker questions and a general presentation of the project and its goals, then moves to more specific questions to foster an open discussion between the requirements engineer and the interviewee.

5.1.1. Participants

Figure 4 represents the stakeholders identified with respect to the three entry points of the MATERIALIST framework. The rest of this section details the stakeholders and the participants of the elicitation study.

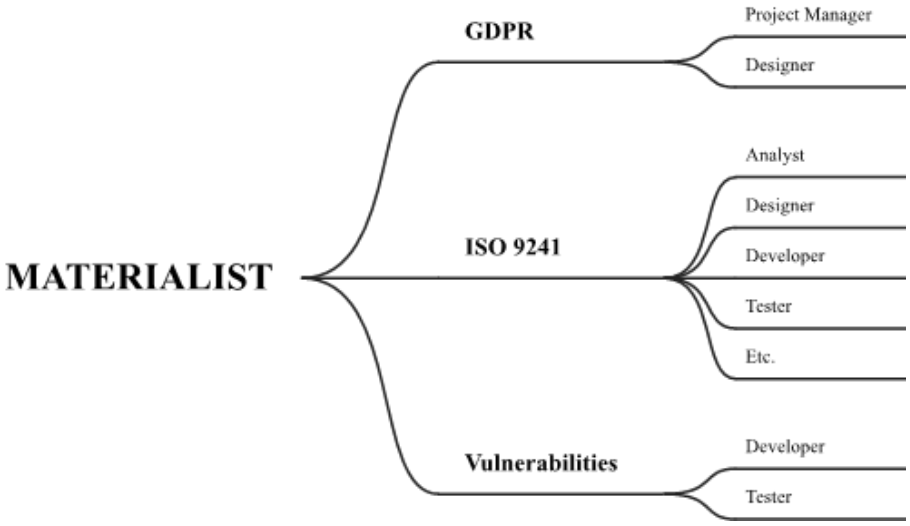


Figure 4. MATERIALIST's stakeholders, identified with respect to the three entry points

The Project Manager (PM) manages the software development project by supervising, monitoring, and guaranteeing that the product is delivered on time and on budget and that the software satisfies the customer's expectations. The Analyst's responsibility is to communicate with the customer to define system requirements by studying the application domain and its procedures. The Designer/Solution Architect is tasked with developing a system architecture that meets the given requirements, selecting the best solutions and technologies in terms of both hardware and software. The Developer is in charge of implementing software based on the specifications set by the analyst and designer, as well as documenting the code. The Tester is responsible for developing and carrying out relevant test cases to

ensure the proper implementation of the program in its many components and compliance with requirements, as well as identifying potential mistakes and problems and their causes.

A total of 9 participants (8 males and 1 female, aged from 24 to 43, M = 33.67, SD = 6.26), including 3 project managers, 1 analyst, and 5 developers, were interviewed. The proficiency level in the security and privacy field of the interviewees was surveyed using a 5-point Likert scale (1 = No experience in this area - 5 = Extensive experience in this area) and resulted in a mean = 3,37 and a standard deviation = 1,19.

5.1.2. Procedure

According to the defined interview protocol, 9 interviews were conducted with potential stakeholders of the platform. The interviews started with a greeting to welcome each participant. Then, they signed a consent form. Subsequently, a concise overview of the interview's goals was provided. Once at ease, the participant received a brief introduction to GDPR, security vulnerabilities, and ISO 9241. The aim was to ensure a common understanding of these concepts. For each of the three entry points, participants were asked to envision specific scenarios, and a series of follow-up questions were posed to delve deeper into their perspectives. Finally, participants were thanked for their time and dismissed. The interviews lasted 30 to 45 mins each, and the answers to the questions were audio-recorded.

5.1.3. Analysis and Results

To extract relevant requirements to be considered during the design and development of the tool, a thematic analysis (TA) has been conducted. TA is a methodology used to identify and extract patterns (called *themes)* within qualitative data [12]. Specifically, this methodology consists of 6 steps: familiarizing with the data, generating codes, mapping together codes that share similar meanings, reviewing potential themes, defining and naming themes, and producing a report. Familiarizing with the data involves identifying items of interest and preparing a baseline for subsequent steps. Generating codes involves defining labels that describe the interesting aspects of the data and highlighting relevant sentences in the transcripts. These short descriptions capture the essence and meaning of the highlighted data. Mapping between codes and themes helps identify the most relevant and recurring codes that better describe the data. Reviewing potential themes helps refine and understand if the identified themes provide a proper and accurate description of the data. Each theme should be accompanied by a short description that defines its meaning. The report then describes the analytic commentary, data extracts, and defined themes.

A series of themes were created following this methodology in order to identify relevant requirements directly from possible platform users to better design and shape the tool in response to the actual demands of its stakeholders.

Theme 1 – **questionnaire**: it is the most recurring theme among the interviewees. Five stakeholders expressed their concerns about not being able to ultimately select the PDPs useful to their needs due to their limited knowledge and the recognized complexity and criticality of security and privacy. A questionnaire to be proposed to the users at the beginning of the process regarding the applicative context and the envisaged architecture of the solution seemed to be the most common solution to kick-start the selection of the patterns supporting stakeholders' jobs.

Theme 2 **– compare PDP according to metrics**: given the extensive number of PDPs and the complementarity of some of them, concerns were raised about being able to successfully choose between similar patterns. To this end, a comparison using metrics was suggested to support the choice of PDPs.

Theme 3 – **group GDPR under principles/rights**: among the three entry points of the framework, the GDPR was recognized, as expected, as the most complex one, being the furthest from the technological point of view of the interviewees. Different interviewees expressed their concerns in actually trying to understand the meaning of the various requirements under the regulation articles. To overcome this

issue, it was proposed to further group the GDPR under relevant principles and rights to help the navigation of the knowledge base and the understandability of the privacy requirements.

Theme 4 – **templates**: to facilitate their job, 2 participants (one developer and one PM) expressed the need to kick-start the process by having available some project templates to start from.

Theme 5 – **search bar**: given the extensive number of entities in MATERIALIST, an interviewee (developer) proposed a search bar to filter the PDPs or the GDPR articles.

Theme 6 – **group PDPs under functionalities**: given the extensive number of PDPs in MATERIALIST, an interviewee (developer) proposed grouping PDPs into categories depending on their related functionality.

Theme 7 – **show PDPs applicative examples**: given the complexity of some PDPs, an interviewee (developer) suggested highlighting real examples of the PDPs to ease developers' implementation process.

Theme 8 – **navigate the knowledge base**: a participant (developer) expressed the need to freely navigate the knowledge base of the system so that the tool can also be used as a reference independently from the other offered functionalities.

Theme 9 – **need for a legal expert**: an interviewee (PM) also proposed to consider as part of the process another user in the role of the legal expert.

Theme 10 – **stop and resume the process (questionnaire)**: an interviewee (analyst) proposed the possibility to stop and resume the process of the questionnaire at any time to visualize the intermediate results and have a non-definitive overview of the project.

5.2. Implementation of the platform

The following sections summarize the MATERIALIST web platform's implementation details. The platform was built with the JHipster framework, with the front-end using React Redux and the back-end using Java Spring Boot. The web app is deployed using Docker Compose, using MariaDB to manage the database. MATERIALIST can be found at https://materialist.ddns.net

5.2.1. Data model

According to the mapping already presented in Section 4, as well as the result of the elicitation study, an Entity-Relationship (ER) diagram has been defined covering the different entities of the framework and their relationships (see Figure 5). This diagram mainly focuses on the mapping from Vulnerabilities, GDPR Articles, and ISO 9241-210 to PDPs.

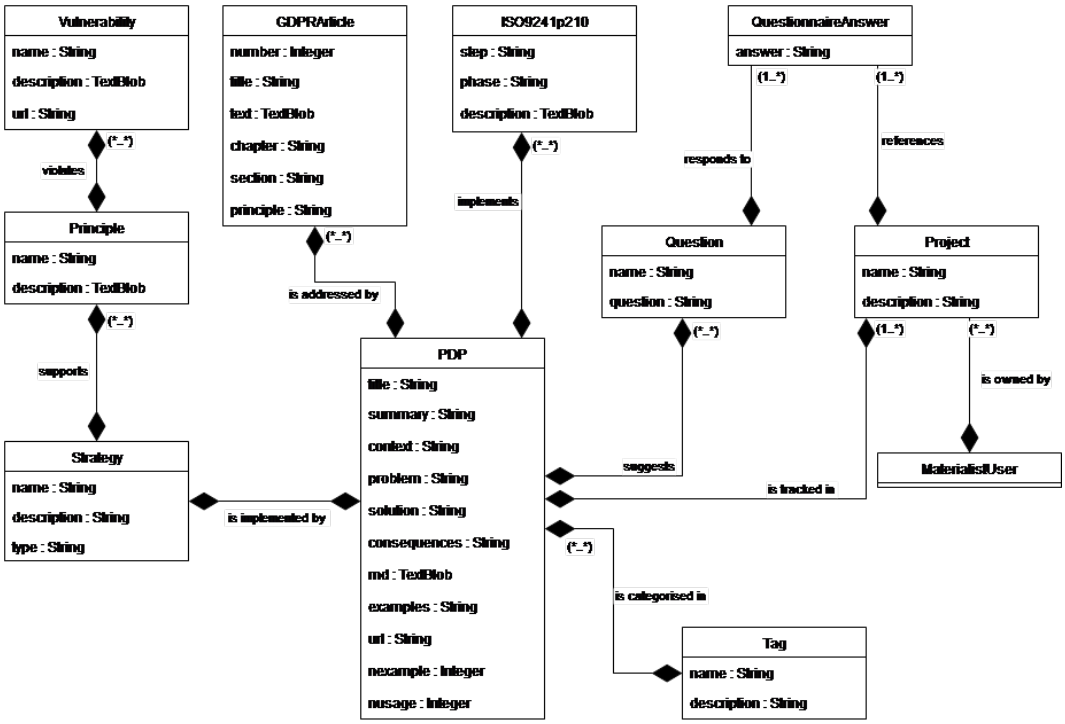


Figure 5 MATERIALIST Knowledge-Base ER diagram

An entity *Project* represents the system under investigation, which is linked to PDPs based on the user's preference. Furthermore, as a result of the elicitation study, new qualities and entities are created to reflect the recommendations of the various stakeholders.

The *question* and *questionnaireAnswer* entities have been defined to link the different questions of the questionnaire to the PDPs to be suggested according to their answers (Theme 1). These entities also serve to store the current state to support the resuming and review of the questionnaire (Theme 10).

The *nexample* and *nusage* attributes of the PDP entity have been introduced to provide metrics to support the comparison between PDPs, suggesting patterns that are more commonly used and more comprehensive (Theme 2). Specifically, the *nexample* attribute represents the number of applicative examples available for a specific pattern, while the *nusage* attribute represents the number of times that specific pattern has been added for implementation in a project.

The *chapter* and the *section* attributes of the *GDPRArticle* entity have been added to provide an easier categorization and navigation between the different articles of the GDPR (Theme 3).

The *tag* entity has been introduced to group PDPs for easier navigation (Theme 6).

The *materialistUser* entity supports introducing different user roles in the platform (eventually responding to Theme 9).

5.2.2. User interface

The interaction with the platform has been divided into two primary sections. The *Knowledge Base* section allows users to investigate the framework's various entry points and related mappings (in accordance with Theme 8). Instead, the *Project* section allows users to select and track the PDPs to be implemented, providing an overview of each phase of the ISO 9241-210 lifespan and the GDPR articles to which the PDPs answer. Each user can access a sample project as an introduction to the platform's functionality (Theme 4).

When starting a new project, the user is prompted to complete a questionnaire to be assisted in selecting the PDPs that should be implemented based on the software requirements (see Figure 6).

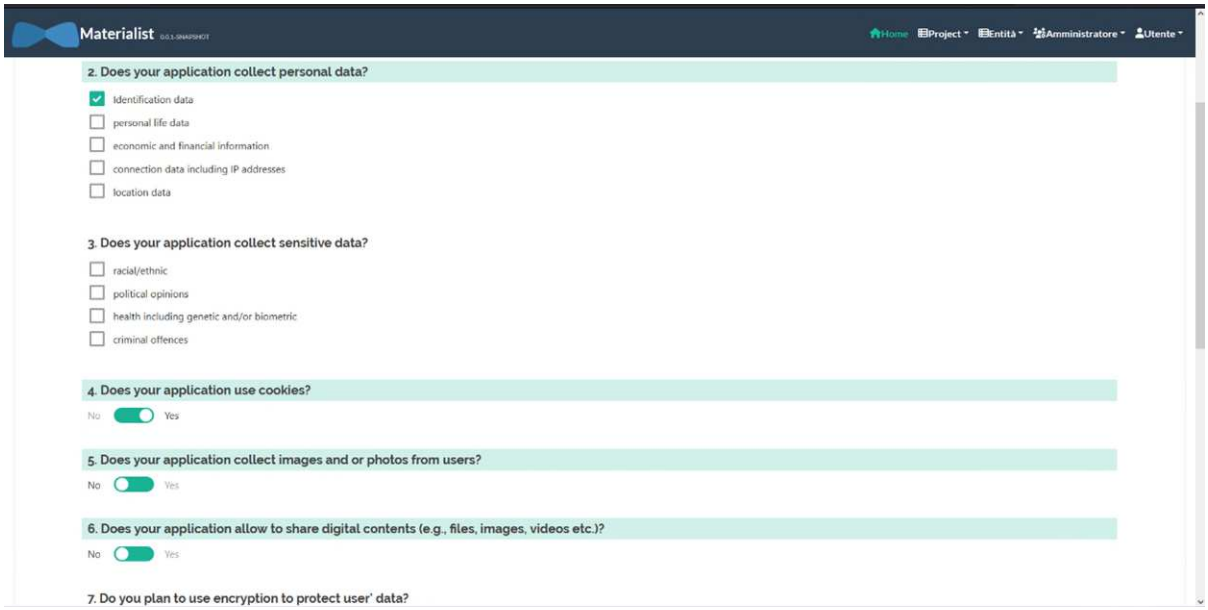


Figure 6 MATERIALIST Questionnaire

After completing the questionnaire, the user is presented with an overview page (see Figure 7) where they may view the various PDPs monitored in the project as well as their implementation details, including an application sample of the patterns (Theme 7). An info tooltip tells the user why a PDP was added to the project (for example, because they correctly answered a question or it was manually added).

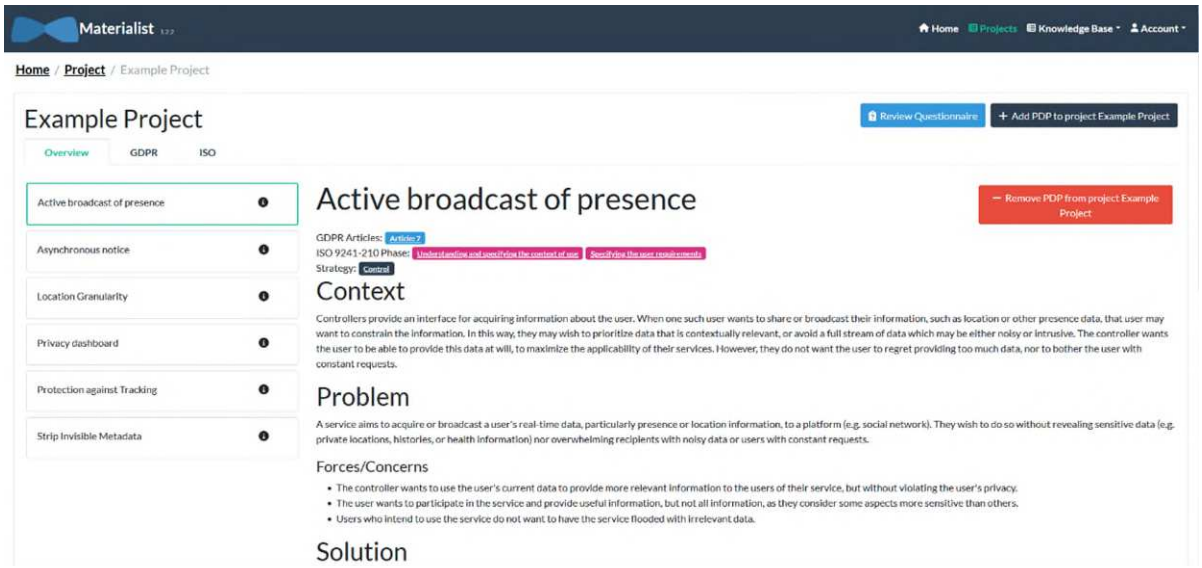


Figure 7 MATERIALIST Project Overview

A "GDPR" tab reports which GDPR articles are covered by which PDPs' implementation, whilst an "ISO" tab indicates which phase each PDP should be evaluated and executed in. The user can add (or remove) any PDP from the project by searching the knowledge base for patterns that match the system requirements based on the various mappings (Theme 5).

6. Exploratory evaluation of the MATERIALIST web platform

An exploratory user study was conducted to assess the platform's usability and utility and to verify whether MATERIALIST could effectively support stakeholders in designing and building GDPR-compliant software. Eight volunteers (8 males aged 23 to 39, M = 30,4, SD = 5,5), different from the

ones recruited in the previous study, were involved. Their security and privacy expertise were evaluated using a 5-point Likert scale, with a mean of 3.22 and a standard deviation of 0.97.

6.1. Measurements and instruments

Five of the eight study sessions were held in an office environment familiar to the participants and on their own PCs for added convenience. The remaining three sessions were conducted remotely via Microsoft Teams, with participants sharing their screens with the facilitator. All participants were potential stakeholders in the platform, ranging from junior to senior software developers (6), analysts (1) and project managers (1).

Quantitative and qualitative data were collected. Regarding quantitative data, task success rate and task time were measured to determine the effectiveness and efficiency of the participants in completing the given tasks. In addition, the Italian-translated USE (Usefulness, Satisfaction, and Ease of Use) questionnaire, a standardized assessment tool for measuring the usability and user experience of a product, was administered [35]. It consists of 30 statements related to usefulness, ease of use, ease of learning and satisfaction; participants are asked to rate their agreement with each statement on a 7-point Likert scale. For qualitative data, observational notes were taken during the study, and a second questionnaire with two open-ended questions about the pros and cons of the platform was administered.

6.2. Procedure

Each session followed the same format. The conductor began by briefly presenting the study and asking to sign a consent form; all participants agreed. Then, the conductor provided a quick overview of the platform, outlining its objective, the structure on which it is built, and its important features. The participants were then instructed to perform the tasks using the Thinking Aloud methodology. Each participant completed the tasks listed below in the order given by the study's objective:

- T1: Find in the example project the pattern related to invisible metadata;
- T2: Find in the example project the pattern to be implemented in the ISO phase relating to the specification of the context of use;
- T3: Find the GDPR article corresponding to the pattern related to impact assessment;
- T4: Answer the questionnaire for creating a new project;
- T5: Add a pattern of the demonstrated strategy to the new project;

The two questionnaires were administered after the execution of the tasks. During the entire procedure, an observer took notes about participants' comments and interactions.

6.3. Data analysis and results

The analysis of the data collected during the study starts by considering the task success rate (details in Table 3) and the execution times (details in Table 4). Tasks were considered failed if the final goal was not successfully reached, or if participants took a path they were not supposed to take. In terms of time, we considered different thresholds towards task success, as each task presented a different level of difficulty, specifically: Task 1 > 60s, Task 2 > 45s, Task 3 > 120s; Task 4 > 200s; and Task 5 > 150s.

Task 1 and Task 2 were executed with fewer errors (both had 87,5% success rates), and participants were quick in completing them since they spent respectively 32 sec and 25 sec. Analyzing the video of Task 1 and Task 2 failed by participants, it emerged that P8 failed Task 1 since it followed the wrong path to reach the goal, missing the PDPs directly listed in the Overview tab, while P3 failed Task 2 since they clicked on the wrong ISO phase, causing a lost in translation issue.

In contrast, Task 3 and Task 5 had the lowest success rate (75%) and participants were quite slow in completing them. In Task 3 both participants who failed missed the search bar to find the Federated Privacy Impact Assessment PDP; on the contrary, participants who immediately noticed the search bar quickly completed the task. For Task 5, one participant (P4) failed the task declaring that it was quite

difficult, while P8 missed clicking the button to add the PDP to the project. In any case, this task was quite troublesome even for the other participants, who took a long time to reach the goal. This task revealed that most participants would have preferred to search and add a PDP to a project directly from the PDP catalog page rather than going to the Knowledge Base section looking for the demonstrated strategy as the task required them to do. Moreover, two participants (P4 and P7) first tried to use the "Edit" button to add a new PDP to the project, revealing the misleading purpose of that button.

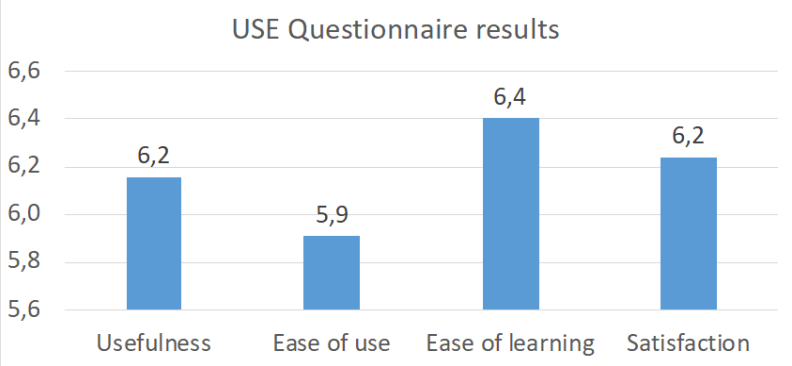


Figure 8. Results of the four dimensions (usefulness, ease of use, ease of learning, and satisfaction) of the USE questionnaire.

The analysis of the USE questionnaire has been conducted along its four dimensions, i.e., usefulness, ease of use, ease of learning, and satisfaction. Regarding *usefulness* (see Figure 8), it resulted in an average score of 6.2 out of 7, which indicates that users find the web platform to be very useful. This is a positive result, showing that the platform is meeting users' needs effectively. Looking in detail at the eight questions of this dimension (see Figure 9), it is evident that Q5 (It makes the things I want to accomplish easier to get done) and Q7 (It meets my needs) achieved the lowest score, despite still obtaining positive results; these scores indicate that we need to make it easier to accomplish tasks and that we need to meet user needs slightly better. On the contrary, Q3 (It is useful) highlights that users perceived the platform as very useful in their work.

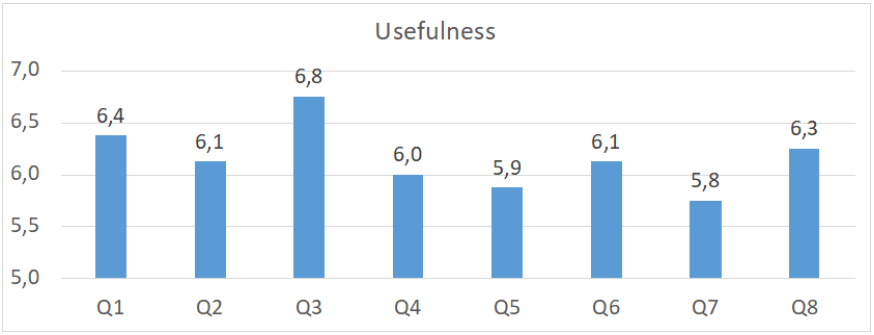


Figure 9. Detailed results of the "usefulness" dimension of the USE questionnaire

In terms of *ease of use*, the platform is considered fairly easy to use, with a score of 5.9 out of 7. Although this is not a perfect score, it is still high and suggests that users generally find the platform easy to use, while there may be some room for improvement in terms of usability. As shown in Figure 10, the questionnaire items that most negatively affected this score are Q15 (I can use it without written instructions), Q18 (I can recover from mistakes quickly and easily), and Q12 (It requires the fewest possible steps to accomplish what I want to do with it). They clearly indicate some aspects that can be optimized to improve usability, such as the possibility of quickly training users before using the platform (Q15), improving the ways users can recover from possible mistakes (Q18), and reducing the number of steps to perform some actions, for example by adding a shortcut for the most used function (Q12).

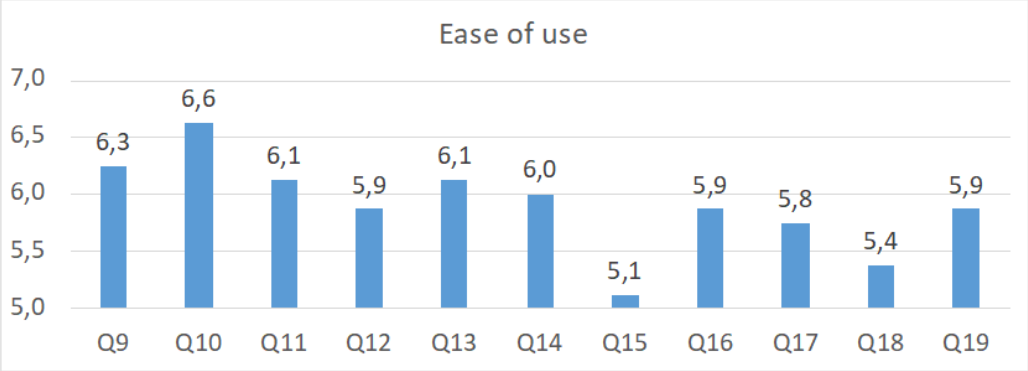


Figure 10. Detailed results of the "ease of use" dimension of the USE questionnaire

Regarding the *ease of learning*, the achieved score of 6.4 out of 7 suggests that users find the platform very easy to learn, indicating that users can quickly grasp how to use the platform and that there is a low learning curve. The details reported in Figure 11 clearly show that all questionnaire items in this dimension contributed in a fairly equal way to the final score, with no particular positive or negative choices.

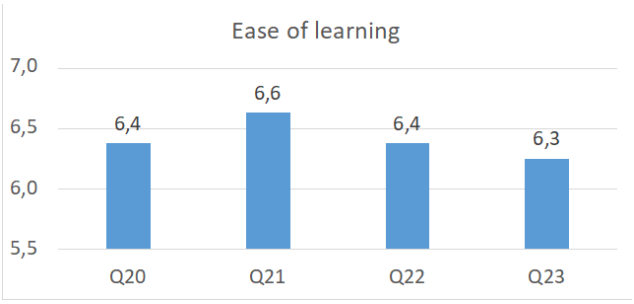


Figure 11. Detailed results of the "ease of learning" dimension of the USE questionnaire

Finally, the satisfaction dimension, with a score of 6.2 out of 7, indicates a very high level of satisfaction among users, which is a positive sign that they are satisfied with their overall experience. Only three items in this dimension might deserve attention, i.e., Q26 (It is fun to use), Q28 (It is wonderful), and Q29 (I feel I must have it). The details depicted in Figure 12 indicate that, while Q26 and Q28 can be associated with the hedonic aspect of the platform, which is not critical to its purpose, more interesting evidence may come from Q29, as it reveals the need to better adapt the platform to the user's needs.

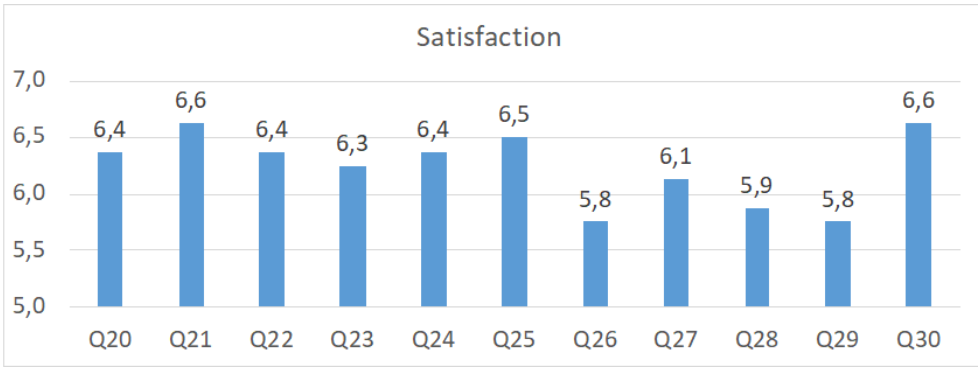


Figure 12. Detailed results of the "satisfaction" dimension of the USE questionnaire

Regarding the qualitative data, we analyzed the observer notes and participants' answers to the open questions to identify the most important cons (Table 5) and pros (Table 6) of the platform. Regarding the cons, P3 and P4 complained that the search bar was not immediately visible. P4 and P7 also

suggested altering the color palette, perhaps to accommodate the dark mode. Three other participants (P2, P6, and P8) expressed how much they would have liked to have seen more details and explanations regarding the various entities and concepts included in the tool.

In contrast, the search functionality and the first questionnaire were highly valued in relation to the three most positive characteristics (e.g., P2, P6, P8). Numerous participants described the platform as intuitive, easy to use, and friendly (e.g., P1, P3, P4, P5, P7).

6.4. Suggested improvements

Our study demonstrated that the MATERIALIST platform is generally perceived as effective and user-friendly. However, the study allowed us to shed light on some limitations and areas for improvement. Regardless of the platform's discovered positive attributes, some issues persist. Notably, participants encountered difficulties when attempting to navigate the platform without written instructions. In particular, less experienced users struggled more. To address these limitations and enhance the overall user experience, it is imperative to streamline the management of PDPs and improve the intuitiveness and visibility of the platform's features. Based on the insights gained from the user tests, the next iteration of platform development will incorporate the following enhancements: introduction of a concise tutorial outlining the platform's primary features, providing more contextualized descriptions clarifying their purpose; enhancement of the visibility and capabilities of the search function; and a streamlining of the process for adding PDPs to a project, making it more efficient. By implementing these enhancements, we aim to address the identified limitations, providing users with an even more effective and intuitive platform for selecting PDPs.

7. Conclusion and future work

This article explored *how software development can be supported to ensure GDPR compliance*, which is a critical and under-supported aspect in the current state of the art. To this aim, we presented MATERIALIST, a web platform designed to support user-centric, secure, and privacy-aware software engineering practices, catering to both forward and backward processes. MATERIALIST empowers users to efficiently select Privacy Design Patterns (PDPs), offering multiple entry points, i.e., GDPR articles, code vulnerabilities, and ISO 9241-210 phases.

To enhance the PDP selection process, as future work we plan to work on defining intermediate layers that provide guidance for any initial entry point; for example, scenarios, functional requirements and project features will be part of these layers to facilitate the PDPs selection. Additionally, we are delving into heuristic and metric-based approaches, leveraging artificial intelligence, to offer context-driven recommendations for the most suitable PDPs, enhancing the platform's effectiveness. Furthermore, we plan to further expand the PDPs catalog by defining architectural patterns and User Interface patterns for each one. This expansion aims to provide support to designers and developers throughout various software development activities, such as coding and design. We believe that by providing enhanced guided access and an expanded catalog of PDPs, MATERIALIST has the potential to evolve into a valuable asset within the domains of usable security and privacy-conscious software engineering. Finally, a broader user study in the form of a longitudinal study will be conducted in more realistic settings, such as enterprise projects, to better evaluate the usability, acceptability and effectiveness of the MATERIALIST framework.

Acknowledgments

This work is partially supported by the Italian Ministry of University and Research (MUR) under grant PRIN 2022 PNRR "DAMOCLES: Detection And Mitigation Of Cyber attacks that exploit human vuLnerabilitiES" – CUP: H53D23008140001.

This work is partially supported by the co-funding of the European Union - Next Generation EU: NRRP Initiative, Mission 4, Component 2, Investment 1.3 – Partnerships extended to universities, research centres, companies and research D.D. MUR n. 341 del 5.03.2022 – Next Generation EU (PE0000014 – "Security and Rights In the CyberSpace – SERICS" - CUP: H93C22000620001).

The research of Andrea Esposito is funded by a Ph.D. fellowship within the framework of the Italian "D.M. n. 352, April 9, 2022"- under the National Recovery and Resilience Plan, Mission 4, Component 2, Investment 3.3 – Ph.D. Project "Human-Centered Artificial Intelligence (HCAI) techniques for supporting end users interacting with AI systems", co-supported by "Eusoft S.r.l." (CUP H91I22000410007).

The research of Francesco Greco is funded by a PhD fellowship within the framework of the Italian "D.M. n. 352, April 9, 2022"- under the National Recovery and Resilience Plan, Mission 4, Component 2, Investment 3.3 - PhD Project "Investigating XAI techniques to help user defend from phishing attacks", co-supported by "Auriga S.p.A." (CUP H91I22000410007).

Declarations

Competing Interests. The authors declare no competing interests.

Conflict of interest. The authors declare that they have no conflict of interests.

Authors' contribution. Conceptualization: Marco Saltarella, Giuseppe Desolda, Rosa Lanzilotti; Investigation: Marco Saltarella, Giuseppe Desolda, Rosa Lanzilotti, Andrea Esposito, Francesco Greco; Methodology: Marco Saltarella Giuseppe Desolda, Rosa Lanzilotti; Resources: Marco Saltarella; Software: Marco Saltarella; Supervision: Rosa Lanzilotti, Giuseppe Desolda; Writing – original draft: Marco Saltarella, Giuseppe Desolda; Writing – review & editing: Giuseppe Desolda, Andrea Esposito, Francesco Greco, Rosa Lanzilotti.

Data availability. All data are available upon request from the corresponding author.

References

1. Ackerman MS, Mainwaring SD (2005) Privacy Issues and Human-Computer Interaction. In: Cranor LF, Garfinkel S (eds) Security and usability: designing secure systems that people can use. O'Reilly Media, Cambridge, MA, pp 19–26

2. Acquisti A, Brandimarte L, Loewenstein G (2020) Secrets and Likes: The Drive for Privacy and the Difficulty of Achieving It in the Digital Age. J Consum Psychol 30:736–758. doi: 10.1002/jcpy.1191

3. Adam Satariano Google Is Fined $57 Million Under Europe's Data Privacy Law. In: New York Times. https://www.nytimes.com/2019/01/21/technology/google-europe-gdpr-fine.html

4. Alhazmi A, Arachchilage NAG (2021) I'm all ears! Listening to software developers on putting GDPR principles into software development practice. Personal and Ubiquitous Computing 25:879–892. doi: 10.1007/s00779-021-01544-1

5. Alkubaisy D, Piras L, Al-Obeidallah MG, Cox K, Mouratidis H (2021) A Framework for Privacy and Security Requirements Analysis and Conflict Resolution for Supporting GDPR Compliance Through Privacy-by-Design. In: International Conference on Evaluation of Novel Approaches to Software Engineering. Springer, pp 67–87

6. Altman I (1975) The environment and social behavior: privacy, personal space, territory, crowding. Brooks/Cole Pub. Co, Monterey, Calif

7. Assal H, Hurtado S, Imran A, Chiasson S (2015) What's the deal with privacy apps?: a comprehensive exploration of user perception and usability. In: Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia. ACM, Linz Austria, pp 25–36

8. Baldassarre MT, Barletta VS, Caivano D, Piccinno A (2020) A Visual Tool for Supporting Decision-Making in Privacy Oriented Software Development. In: Proceedings of the International Conference on Advanced Visual Interfaces. Association for Computing Machinery, New York, NY, USA

9. Baldassarre MT, Barletta VS, Caivano D, Scalera M (2019) Privacy Oriented Software Development. Communications in Computer and Information Science 1010:18–32. doi: 10.1007/978-3-030-29238-6_2

10. Baldassarre MT, Barletta VS, Caivano D, Scalera M (2020) Integrating security and privacy in software development. Software Qual J 28:987–1018. doi: 10.1007/s11219-020-09501-6

11. Borchers JO (2000) A pattern approach to interaction design. In: Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques. ACM, New York City New York USA, pp 369–378

12. Braun V, Clarke V (2012) Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ (eds) APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological. American Psychological Association, Washington, pp 57–71

13. Brodin M (2019) A Framework for GDPR Compliance for Small- and Medium-Sized Enterprises. European Journal for Security Research 4:243–264. doi: 10.1007/s41125-019-00042-z

14. Capodieci A, Mainetti L (2019) Business Process Awareness to Support GDPR Compliance. In: Proceedings of the 9th International Conference on Information Systems and Technologies. Association for Computing Machinery, New York, NY, USA

15. Carly Page EU hits Amazon with record-breaking $887M GDPR fine over data misuse. In: Tech Crunch. https://techcrunch.com/2021/07/30/eu-hits-amazon-with-record-breaking-887m-gdpr-fine-over-data-misuse/

16. Cavoukian A (2009) Privacy by design: The 7 foundational principles

17. Cavoukian A, Chanliau M (2013) Privacy and security by design: A convergence of paradigms. Information and Privacy Commissioner, Ontario

18. Cavoukian A, Dixon M (2013) Privacy and security by design: An enterprise architecture approach. Information and Privacy Commissioner of Ontario, Canada

19. Colesky M, Caiza JC, Del Alamo JM, Hoepman J-H, Martín Y-S (2018) A system of privacy patterns for user control. In: Proceedings of the 33rd Annual ACM Symposium on Applied Computing. pp 1150–1156

20. Colesky M, Hoepman J-H, Hillen C (2016) A critical analysis of privacy design strategies. In: 2016 IEEE Security and Privacy Workshops (SPW). IEEE, pp 33–40

21. Degeling M, Utz C, Lentzsch C, Hosseini H, Schaub F, Holz T (2019) We Value Your Privacy ... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy. Informatik Spektrum 42:345–346. doi: 10.1007/s00287-019-01201-1

22. Drozd O, Kirrane S (2019) I Agree: Customize Your Personal Data Processing with the CoRe User Interface. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 11711 LNCS:17–32. doi: 10.1007/978-3-030-27813-7_2

23. Habib H, Pearman S, Wang J, Zou Y, Acquisti A, Cranor LF, Sadeh N, Schaub F (2020) “It’s a Scavenger Hunt”: Usability of Websites’ Opt-Out and Data Deletion Choices. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp 1–12

24. Hickey AM, Davis AM (2004) A Unified Model of Requirements Elicitation. Journal of Management Information Systems 20:65–84. doi: 10.1080/07421222.2004.11045786

25. Iachello G, Hong J (2007) End-User Privacy in Human-Computer Interaction. FNT in Human-Computer Interaction 1:1–137. doi: 10.1561/1100000004

26. IAPP IAPP-FTI Consulting Privacy Governance Report 2020

27. IBM (2022) Cost of a data breach 2022. https://web.archive.org/web/20220930044228/www.ibm.com/reports/data-breach. Accessed 7 Nov 2022

28. International Telecommunication Union TSS (1991) Security Architecture for Open Systems Interconnection (OSI) for CCITT Applications

29. ISO/IEC 27701 (2019) Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines. International Organization for Standardization, Geneva, CH

30. ISO/TC 159/SC 4 (2019) ISO 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems

31. Italian Data Protection Authority Act 9870832 of March 30, 2023

32. Jess Weatherbed OpenAI’s regulatory troubles are only just beginning. In: The Verge. https://www.theverge.com/2023/5/5/23709833/openai-chatgpt-gdpr-ai-regulation-europe-eu-italy

33. Johansen J, Fischer-Hübner S (2020) Making GDPR usable: A model to support usability evaluations of privacy. IFIP Advances in Information and Communication Technology 576 LNCS:275–291. doi: 10.1007/978-3-030-42504-3_18

34. Lioudakis GV, Koukovini MN, Papagiannakopoulou EI, Dellas N, Kalaboukas K, de Carvalho RM, Hassani M, Bracciale L, Bianchi G, Juan-Verdejo A, Alexakis S, Gaudino F, Cascone D, Barracano P (2020) Facilitating GDPR Compliance: The H2020 BPR4GDPR Approach. In: Pappas IO, Mikalef P, Dwivedi YK, Jaccheri L, Krogstie J, Mäntymäki M (eds) Digital Transformation for a Sustainable Society in the 21st Century. Springer International Publishing, Cham, pp 72–78

35. Lund AM (2001) Measuring usability with the use questionnaire. Usability interface 8:3–6

36. McParlan J, van der Linden D (2021) Privacy labels should go to the dogs. In: Eight International Conference on Animal-Computer Interaction. ACM, Bloomington IN USA, pp 1–10

37. Mead NR, Stehney T (2005) Security quality requirements engineering (SQUARE) methodology. ACM SIGSOFT Software Engineering Notes 30:1–7. doi: 10.1145/1082983.1083214

38. Pattakou A, Mavroeidi A-G, Diamantopoulou V, Kalloniatis C, Gritzalis S (2018) Towards the Design of Usable Privacy by Design Methodologies. In: 2018 IEEE 5th International Workshop on Evolving Security & Privacy Requirements Engineering (ESPRE). IEEE, Banff, AB, pp 1–8

39. Raschke P, Küpper A, Drozd O, Kirrane S (2018) Designing a GDPR-compliant and usable privacy dashboard. IFIP Advances in Information and Communication Technology 526:221–236. doi: 10.1007/978-3-319-92925-5_14

40. Renaud K, Shepherd LA (2018) How to make privacy policies both GDPR-compliant and usable. In: 2018 International Conference on Cyber Situational Awareness, Data Analytics and Assessment, CyberSA 2018

41. Riva GM, Vasenev A, Zannone N (2020) SoK: Engineering Privacy-Aware High-Tech Systems. In: Proceedings of the 15th International Conference on Availability, Reliability and Security. Association for Computing Machinery, New York, NY, USA

42. Senarath A, Arachchilage NAG (2018) Why developers cannot embed privacy into software systems?: An empirical investigation. In: Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering 2018. ACM, Christchurch New Zealand, pp 211–216

43. Seun Solomon Bakare, Adekunle Oyeyemi Adeniyi, Chidiogo Uzoamaka Akpuokwe, Nkechi Emmanuella Eneh (2024) Data Privacy Laws and Compliance: A Comparative Review of the Eu Gdpr and Usa Regulations. Comput sci IT res j 5:528–543. doi: 10.51594/csitrj.v5i3.859

44. Sharma A, Nilsen TB, Brenna L, Johansen D, Johansen HD (2021) Accountable Human Subject Research Data Processing using Lohpi. In: Proceedings of the ICTeSSH 2021 conference. PubPub, Virtual conference

45. Shneiderman B, Plaisant C, Cohen M, Jacobs S, Elmqvist N, Diakopoulos N (2016) Grand challenges for HCI researchers. interactions 23:24–25. doi: 10.1145/2977645

46. Smirnova Y, Travieso-Morales V (2024) Understanding challenges of GDPR implementation in business enterprises: a systematic literature review. IJLMA 66:326–344. doi: 10.1108/IJLMA-08-2023-0170

47. Sobolewski M, Mazur J, Paliński M (2017) GDPR: A step towards a user-centric internet? Intereconomics 52:207–213. doi: 10.1007/s10272-017-0676-5

48. Tarique M, Ansari J, Abdullah Baz ·, Alhakami · Hosam, Alhakami W, Kumar R, Raees ·, Khan A (2021) P-STORE: Extension of STORE Methodology to Elicit Privacy Requirements. Arabian Journal for Science and Engineering 46:8287–8310. doi: 10.1007/s13369-021-05476-z

49. Westin AF (2013) Privacy and freedom, 2nd edition. International Association of Privacy Professionals, Portsmouth, NH

50. Wong RY, Mulligan DK (2019) Bringing Design to the Privacy Table: Broadening "Design" in "Privacy by Design" Through the Lens of HCI. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp 1–17

51. Zowghi D, Coulin C (2005) Requirements Elicitation: A Survey of Techniques, Approaches, and Tools. In: Aurum A, Wohlin C (eds) Engineering and Managing Software Requirements. Springer-Verlag, Berlin/Heidelberg, pp 19–46

52. (2022) 2022 Data Breach Investigations Report. In: Verizon Business. https://www.verizon.com/business/resources/reports/dbir/. Accessed 7 Nov 2022

Appendix

Table 1. Mapping between GDPR Articles and PDPs available at https://privacypatterns.org/

GDPR Article	Privacy Design Pattern
Article 5 Principles relating to processing of personal data	Ambient Notice
	Appropriate Privacy Feedback
	Appropriate Privacy Icons
	Attribute Based Credentials
	Awareness Feed
	Dynamic Privacy Policy Display
	Icons for Privacy Policies
	Impactful Information and Feedback
	Lawful Consent
	Location Granularity
	Obtaining Explicit Consent
	Platform for Privacy Preferences
	Policy Matching Display
	Privacy Aware Wording
	Privacy Awareness Panel
	Privacy Color Coding
	Privacy dashboard
	Privacy icons
	Privacy Labels
	Privacy Mirrors
	Privacy Policy Display
	Strip Invisible Metadata
	Trust Evaluation of Services Sides
Article 6 Lawfulness of processing	[Support] Selective Disclosure
	Decoupling [content] and location information visibility
	Discouraging blanket strategies
	Enable/Disable Functions
	Informed Implicit Consent
	Lawful Consent
	Sign an Agreement to Solve Lack of Trust on the Use of Private Data Context
Article 7 Conditions for consent	[Support] Selective Disclosure
	Active broadcast of presence
	Decoupling [content] and location information visibility
	Enable/Disable Functions
	Obtaining Explicit Consent
	Sign an Agreement to Solve Lack of Trust on the Use of Private Data Context
Article 12	Abridged Terms and Conditions
	Ambient Notice

Transparent information, communication and modalities for the exercise of the rights of the data subject	Asynchronous notice
	Awareness Feed
	Discouraging blanket strategies
	Dynamic Privacy Policy Display
	Icons for Privacy Policies
	Impactful Information and Feedback
	Increasing awareness of information aggregation
	Informed Consent for Web-based Transactions
	Informed Credential Selection
	Informed Implicit Consent
	Layered Policy Design
	Minimal Information Asymmetry
	Obtaining Explicit Consent
	Platform for Privacy Preferences
	Policy Matching Display
	Privacy Aware Wording
	Privacy Awareness Panel
	Privacy Color Coding
	Privacy icons
	Privacy Labels
	Privacy Mirrors
	Privacy Policy Display
	Privacy-Aware Network Client
	Trust Evaluation of Services Sides
	Who's Listening
Article 13 Information to be provided where personal data are collected from the data subject	Outsourcing [with consent]
Article 15 Right of access by the data subject	Asynchronous notice
	Personal Data Table
	Privacy dashboard
	Privacy Mirrors
Article 20 Right to data portability	Personal Data Table
Article 25 Data protection by design and by default	Buddy List
	Encryption with user-managed keys
	Identity Federation Do Not Track Pattern
	Informed Secure Passwords
	Location Granularity
	Masquerade
	Minimal Information Asymmetry
	Negotiation of Privacy Policy
	Onion Routing
	Personal Data Store
	Preventing mistakes or reducing their impact
	Privacy-Aware Network Client

	Private link
	Protection against Tracking
	Pseudonymous Messaging
	Pseudonymous Identity
	Reasonable Level of Control
	Selective access control
	Sticky Policies
	Strip Invisible Metadata
	User data confinement pattern
Article 26 Joint controller	Obligation Management
Article 28 Processor	Obligation Management
Article 32 Security of processing	Added-noise measurement obfuscation
	Aggregation Gateway
	Anonymity Set
	Anonymous Reputation-based Blacklisting
	Encryption with user-managed keys
	Onion Routing
	Private link
	Protection against Tracking
	Pseudonymous Messaging
	Pseudonymous Identity
	Single Point of Contact
	Trustworthy Privacy Plug-in
	Unusual Activities
	Use of dummies
Article 33 Notification of a personal data breach to the supervisory authority	Data Breach Notification Pattern
Article 34 Communication of a personal data breach to the data subject	Data Breach Notification Pattern
Article 35 Data protection impact assessment	Federated Privacy Impact Assessment

Table 2. Mapping between ISO9241-210 and PDPs

ISO9241-210 Phase	**Privacy Design Pattern**
Initial Planning	Incentivized Participation
	Privacy Awareness Panel
Understanding and specifying the context of use	Abridged Terms and Conditions
	Active broadcast of presence
	Anonymous Reputation-based Blacklisting
	Buddy List
	Icons for Privacy Policies
	Impactful Information and Feedback
	Incentivized Participation

	Increasing awareness of information aggregation
	Informed Secure Passwords
	Minimal Information Asymmetry
	Negotiation of Privacy Policy
	Obligation Management
	Obtaining Explicit Consent
	Outsourcing [with consent]
	Pay Back
	Platform for Privacy Preferences
	Preventing mistakes or reducing their impact
	Privacy Aware Wording
	Privacy Awareness Panel
	Privacy Mirrors
	Reciprocity
	Sign an Agreement to Solve Lack of Trust on the Use of Private Data Context
	Trust Evaluation of Services Sides
Specifying the user requirements	[Support] Selective Disclosure
	Abridged Terms and Conditions
	Active broadcast of presence
	Ambient Notice
	Appropriate Privacy Feedback
	Appropriate Privacy Icons
	Asynchronous notice
	Awareness Feed
	Data Breach Notification Pattern
	Decoupling [content] and location information visibility
	Discouraging blanket strategies
	Dynamic Privacy Policy Display
	Enable/Disable Functions
	Encryption with user-managed keys
	Icons for Privacy Policies
	Informed Credential Selection
	Informed Implicit Consent
	Lawful Consent
	Masquerade
	Obtaining Explicit Consent
	Outsourcing [with consent]
	Personal Data Table
	Policy Matching Display
	Privacy Color Coding
	Privacy dashboard
	Privacy icons
	Privacy Policy Display
	Private link

	Reasonable Level of Control
	Selective access control
	Unusual Activities
	Who's Listening
Producing design solutions	[Support] Selective Disclosure
	Added-noise measurement obfuscation
	Aggregation Gateway
	Anonymity Set
	Anonymous Reputation-based Blacklisting
	Appropriate Privacy Icons
	Attribute Based Credentials
	Data Breach Notification Pattern
	Encryption with user-managed keys
	Identity Federation Do Not Track Pattern
	Informed Consent for Web-based Transactions
	Lawful Consent
	Layered Policy Design
	Location Granularity
	Masquerade
	Negotiation of Privacy Policy
	Onion Routing
	Personal Data Store
	Privacy Aware Wording
	Privacy Color Coding
	Privacy dashboard
	Privacy Labels
	Privacy-Aware Network Client
	Private link
	Protection against Tracking
	Pseudonymous Messaging
	Pseudonymous Identity
	Single Point of Contact
	Sticky Policies
	Strip Invisible Metadata
	Trustworthy Privacy Plug-in
	Unusual Activities
	Use of dummies
	User data confinement pattern
Evaluating the design	Federated Privacy Impact Assessment
	Negotiation of Privacy Policy

Table 3. Task success rates

Participant	T1	T2	T3	T4	T5	Average participant success rate
P1	✓	✓	✓	✓	✓	100%
P2	✓	✓	✓	✓	✓	100%
P3	✓	✗	✓	✓	✓	80%
P4	✓	✓	✓	✓	✗	80%
P5	✓	✓	✗	✓	✓	80%
P6	✓	✓	✗	✓	✓	80%
P7	✓	✓	✓	✓	✓	100%
P8	✗	✓	✓	✓	✗	60%
Average success rate	87,5%	87,5%	75%	100%	75%	
Total average success rate						87,2%

Table 4. Tasks execution times

Participant	T1	T2	T3	T4	T5
P1	20s	20s	82s	71s	90s
P2	37s	36s	71s	154s	54s
P3	33s	20s	86s	66s	105s
P4	38s	25s	105s	122s	180s
P5	24s	24s	153s	70s	61s
P6	43s	24s	142s	94s	103s
P7	26s	20s	61s	99s	55s
P8	32s	30s	68s	185s	132s
Average execution time	32s	25s	96s	108s	98s

Table 5. Most 3 negative aspects of the platform

Participant (Role)	Most 3 negative aspects
P1 (Developer)	1) It would be helpful to have a small tutorial when the user first logs in.
P2 (Developer)	-
P3 (Developer)	1) Search bar is not visible at first glance 2) Descriptive info/category related to each item in Project -> GDPR tab
P4 (Analyst)	1) Search bar is not clearly visible 2) Search doesn't search on all fields and doesn't give suggestions when searching 3)The edit button on the project is not clear
P5 (Developer)	1) I recommend a different font/color combination
P6 (Developer)	1) A brief introductory tutorial would be helpful to orient to the very first use. 2) It would be helpful to be able to search for "Strategies" directly from the search bar within the project, and then add the associated PDP. 3) Whenever possible, the information on the "info" icon should also include a brief description inherent to the item.
P7 (Developer)	1) The "save" appears darker and therefore less visible at a glance with respect to the back button. 2) Dark mode support
P8 (Project Manager)	1) I would like to be more guided and have more information about the concepts introduced in the tool (strategy, principle, iso…) 2) ISO: I would call it "phases," and compare it to other frameworks, e.g., PMBOK (waterfall? Agile?) 3) Better highlight the concept of strategies, principles, and vulnerabilities

Table 6. Most 3 positive aspects of the platform

Participant (Role)	**Most 3 positive aspects**
P1 (Developer)	1) Easy to use. 2) Clearly exposes design patterns to be implemented.
P2 (Developer)	1) Useful for the huge amount of data inside 2) Search bar
P3 (Developer)	1) Very intuitive navigation between the various sections 2) It greatly simplifies the design in case of introducing features that must comply with GDPR articles 3) Allows for complete control over the solutions adopted in different projects
P4 (Analyst)	1) Very responsive 2) Well organized 3) easy to understand
P5 (Developer)	1) Ease of use, flexibility. Good job!
P6 (Developer)	1) I find it very useful to use the questionnaire to guide those who are not familiar with GDPR and Privacy Design Patterns. 2) It is very convenient to be able to organize the different designs. 3) The tool is easy to use and quite intuitive.
P7 (Developer)	1) Outstanding graphics, very user friendly, fast and intuitive.
P8 (Project Manager)	1) Wizard 2) Search bar 3) Connection between pattern and GDPR articles