Abstract—The Privacy by Design approach to systems engineering introduces privacy requirements in the early stages of development, instead of patching up a built system afterwards. However, 'vague', 'disconnected from technology', or 'aspirational' are some terms employed nowadays to refer to the privacy principles which must lead the development process. Although privacy has become a first-class citizen in the realm of non-functional requirements and some methodological frameworks help developers by providing design guidance, software engineers often miss a solid reference detailing which specific, technical requirements they must abide by, and a systematic methodology to follow. In this position paper, we look into a domain that has already successfully tackled these problems -web accessibility-, and propose translating their findings into the realm of privacy requirements engineering, analyzing as well the gaps not yet covered by current privacy initiatives.

Index Terms—Privacy Requirements, Requirement decomposition, Privacy principles, Privacy patterns, Privacy by Design

I. INTRODUCTION

The term Privacy by Design (PbD) designates a software design approach that incorporates privacy requirements from the beginning and throughout the software development process, instead of considering them as an afterthought. The term was originally incepted by the Information and Privacy Commissioner of Ontario, Canada, Ann Cavoukian, whose definition includes a set of 7 foundational principles [1] that should guide the development of systems to mitigate privacy concerns from the very beginning. PbD has rapidly become a hot topic in the privacy realm. In 2010, it was recognized as an essential component of privacy and data protection by Privacy Commissioners from around the world [2], and has been introduced in recommended practices and regulations since then by both the US and the EU, among others.

However, engineering Privacy by Design has demonstrated to be a tough task. Most criticisms come from the technical side, arguing that Cavoukian’s principles are “disconnected from technology” [3], “remain vague and leave many open questions about their applications when engineering systems” [4], and “are more aspirational than practical or operational” [5]. Besides, although Cavoukian and her team have published several documents to operationalize PbD [6] in some specific contexts, “she makes little effort to systematize or even summarize the design principles found” [5]. Further, legal or normative privacy principles are also far from technology, and moving them into technical requirements is not an easy job. It practitioners willing to implement PbD cannot find a structured corpus of requirements, and get instead often lost in questions such as what they should do, or how they should do it. Cavoukian herself has recognized that “The next stage of PbD evolution is to translate its "7 Foundational Principles" into more prescriptive requirements, specifications, standards, best practices, and operational-performance criteria” [7].

From our experience, these problems have already been tackled in other technical domains which show structural similarities to privacy, for example, accessibility -the availability of a product or service to be used by all the people, disregarding the abilities or disabilities they might have. Privacy and accessibility share many commonalities: they are categories of non-functional requirements; more specifically, they are quality-in-use attributes whose determination depends on the specific user and context of use. Besides, they involve complex and interdisciplinary issues: both impact the users’ rights and are thus contemplated by legislation in most countries are by corporate social responsibility frameworks. Therefore, we propose reviewing the current status of practice in accessibility requirements and examining whether some of the initiatives in this realm can be ported to privacy requirements engineering.

II. REQUIREMENTS ENGINEERING IN ACCESSIBILITY

Most of the work on the definition of web accessibility requirements currently revolves around the W3C’s Web Accessibility Initiative (WAI): a multi-stakeholder initiative, developing guidelines, strategies, and resources for web accessibility since 1997. More than 200 people from around the world, representing different industries and stakeholders, are active participants; to which thousands of external commenters should be accrued.

The success of W3C WAI can be measured by the adoption of their guidelines among regulations and policies worldwide: it has set the accessibility standards required in many countries [8], has been endorsed by ISO, and is directly included by reference by the most recent regulations for public procurements in both the US and the EU. That is, fragmented regulations have converged to embrace WAI guidelines as the measuring rod for web accessibility, resulting on legal
homogenization. Besides, the application field of WAI guidelines has been extended beyond its initial scope. Despite initially derived from global principles which could be applicable to any product or service, WAI focused only on the specific field of web technologies. However, its recommendations have now extended to deal with other ICT products and inspire guidelines by operating system vendors.

A. Roles and Requirements

W3C WAI has defined a framework consisting of several components that intervene in providing an accessible experience to the end user. These encompass both human components (the end users themselves and the producers or developers) and technical components, for each of which W3C WAI defines a family of accessibility requirements:

- Web contents, encompassing fully-fledged applications as well, which must abide by the Web Content Accessibility Guidelines (WCAG) to be accessible.
- Authoring tools that mediate between the producers and the web contents.
- User agents that mediate between the web contents and the end users.
- Assistive technologies, which users with disabilities lay over the user agents to decorate them with accessibility-specific functions.
- Evaluation tools that developers (ought to) use to assess conformance with accessibility requirements.
- Technical specifications that define the interactions between all those.

B. Requirement Layers and Conformance

Accessibility requirements specified by WCAG are structured into four layers, each refining the previous one: foundational principles, basic guidelines, testable success criteria, and a rich collection of techniques to meet them [9], [10] in different contexts (Fig. 1). WCAG also establishes a conformance framework around these layers to assess whether and how accessible a system is.

A principle defines the foundations necessary for anyone to access and use a system. Four principles are defined: 1) Perceivable, 2) Operable, 3) Understandable, 4) Robust.

A guideline provides the specific goals that authors should work towards in order to meet a principle. Each principle can be decomposed into a fixed set of guidelines. The guidelines are not testable as they are defined, but refine the principles with specific objectives, and provide a structured framework to better understand the lower layers. WCAG defines 12 guidelines in total, e.g. the Perceivable principle is composed by 4 guidelines, respectively dealing with: 1.1) Text alternatives, 1.2) Time-based media; 1.3) Adaptable content; and 1.4) Distinguishable content.

A success criterion is an observable (by the system user) and measurable item (a checkpoint). For each guideline a set of objectively testable success criteria are provided. Objectivity does not necessarily imply automation: many success criteria may need human judgment for validation. Each success criterion is written as a technology-neutral statement that will be either true or false for a specific web content. For example, regarding the guideline 1.1 (Text alternatives) a success criterion is defined as 1.1.1) “All non-text content that is presented to the user has a text alternative that serves the equivalent purpose, except for the situations listed below [...]”.

A technique is a reliable, implementable way to meet one or more success criteria. Each technique sanctioned by WAI is developed into a short document (usually a few pages), including applicability constraints, a detailed description, usage examples, external references, related techniques, and test procedures. Hundreds of techniques exist, covering a broad variety of cases. Regarding technology, “generic techniques” apply to any system, while “technology-specific techniques” only apply where the respective technology is used (e.g. HTML, Adobe Flash, etc.) Regarding conformance: a “sufficient technique” ensures conformance with the respective success criteria when it is applied (yet it is never necessary, as there may be different techniques to satisfy the same success criterion); an “advisory technique” does not directly provide conformance but helps attaining or going beyond it; finally, if a “failure technique” is not circumvented, some success criteria will be irremissibly failed. And regarding applicability, some techniques can be applied to any content, while others apply only in very specific contexts (e.g. emoticons, talking heads). WAI techniques are deemed as informative (optional), instead of normative (mandatory); however, they have proved to be quite useful as an off-the-shelf resource available to developers, who do not need to come up with their own techniques. Following our previous example, in order to meet the success criterion 1.1.1, technique H37 establishes that, when a short description can serve the same purpose and present the same information as the non-text content, for each HTML img element, the author must specify a short text alternative using the alt attribute.

A test case is an atomic test, usually one that is a partial test for a requirement [11], which defines the applicability criteria (elements of the system where the test is applied), the test procedure, and the expected results. Following on our previous example, a test case could be implemented using XPath expressions to automatically check whether all img elements in an HTML document include an alt attribute with a short text. Test cases are not standardized (hence they are shown in gray in Fig. 1) and they are usually internally defined by specific testing frameworks and tools. Nonetheless, there have been some attempts beyond WAI to provide standardized tests and it is currently an area of work.

Each success criterion has a Level of Conformance (i.e. relevance): A (most relevant), AA, and AAA (not so relevant).
Depending on which success criteria a system satisfies, it may respectively hold one of three levels of conformance: Level A (minimum) implies satisfying all the Level A success criteria, Level AA implies satisfying both Level A and Level AA success criteria, and Level AAA implies satisfying Level AAA success criteria in addition. That is, conformance is defined in a most restrictive fashion: if any part of a system (e.g. a single icon) fails a single success criteria for some level X, the system does not conform to that level X. This has important consequences for the assessment of web accessibility: conformance can only be measured as an ordinal level (A, AA, AAA) and a system cannot be merely “60% accessible”.

III. Systematizing Privacy Requirements Engineering

As we have advanced earlier, our position statement can be summarized as follows: the approach followed by the accessibility community to agree on a structured set of requirements can be translated into the privacy realm, thus defining the path to move from privacy principles to detailed, engineering requirements. This position is based on the similarities between both presented in the introduction. Mirroring the approach followed in the accessibility realm would imply:

- Gathering stakeholders from different communities in an open, neutral forum, encompassing software and service developers, product and service providers, data protection authorities, digital rights associations, government agencies, policy makers, end-user communities, and researchers; where they may agree on a shared set of principles from which further requirements are refined.
- Defining a collection of roles that are involved in the provision of privacy-compliant services and assigning privacy requirements to one or more of these roles.
- Structuring privacy requirements into a series of layers, each refining the previous one into more detailed requirements: principles which define the foundations necessary for a system to be privacy-respectful, guidelines which provide the specific goals that authors should work toward in order to meet a principle, success criteria which define observable, testable and measurable items (checkpoints) for each guideline, and techniques which define a reliable, implementable way to meet several success criteria.
- Providing an evolving catalogue of privacy patterns or techniques, including both technology-neutral and technology-specific patterns, which instruct engineers on the specific ways to meet privacy requirements, who can resort to this catalogue to choose the most appropriate techniques.
- Defining levels of relevance of privacy requirements at the success criteria layer, so that different levels of compliance with the standardized body of requirements can be required, targeted, claimed, and certified.

A. Privacy Principles

The concept of abstract principles is not exotic to the privacy realm. Although different sets of privacy principles have long been defined by different organizations, most of them source from data protection laws, which set the domain rules applicable to systems and software. Among the most relevant, we encounter [12]: the Fair Information Principles originally recorded in 1973 and later expanded in the OECD guidelines, which have inspired much subsequent legislation; the EU Data Protection Directive 95/46/EC and the new General Data Protection Regulation proposal, together with the Safe Harbor principles; the US Federal Trade Commission’s Fair Information Practice Principles; and the Asia-Pacific Economic Cooperation’s Privacy Framework.

Some standardization bodies have started to work on compiling and abstracting a homogeneous set of principles from those above mentioned and others. We should remark the principles of ISO/IEC 29100 [13] and OASIS Privacy Management Reference Model and Methodology (PMRM) [14], based on previous work by the International Security, Trust and Privacy Alliance (ISTPA). These principles are indeed closer to the technical domain e.g. data minimization, use limitation, consent and choice.

On the other hand, Cavoukian’s PbD seven foundational principles provide an intermediate approach: they are focused on the engineering process, rather than on legal requirements; yet they are rather abstract and mix system-oriented principles (e.g. “Privacy as the default setting”, “Visibility and Transparency - Keep it Open”) with design-process-oriented ones (“Privacy Embedded into Design”).

Besides, some authors propose extending the current set of privacy principles, as they argue that they are not enough, either because they only focus on data protection while forgetting other types of privacy [12], or because they focus on procedural aspects while leaving aside structural issues which have a larger impact on consumer rights [15].

As we have seen, there are already several initiatives which state the principles that must be met by privacy-compliant systems. However, there is not yet a consensus on a closed list of principles: the work by ISTPA made a big step in compiling principles from elsewhere, and it has later been picked up by OASIS PMRM, yet there are still divergent works at PbD, ISO, the academia, etc. We are suggesting that a twofold agreement could be assumed, by one of these standardization forums accommodating the privacy community at large, and by the community assuming the results output by that body.

B. Roles for Privacy

The privacy community has recognized the need to identify the roles that participate in privacy-affecting interactions. For example, the OASIS PMRM requires identifying: 1) Participants who have operational privacy responsibilities in any stage of the lifecycle (“creating, managing, interacting with, or otherwise subject to personal information”), 2) Privacy Domains which are subject to the (physical or logical) control of the same owner, and 3) Systems that accomplish a specific function (collect, communicate, process, store or dispose.
personal information) within a privacy domain. This specification does not clearly define roles and their responsibilities, though it establishes the need to identify them.

Other initiatives have gone beyond and defined the roles that usually take part; for example, the UK Data Protection Act and ISO/IEC 29100 identify the following basic roles:

- A Data Subject (or PII Principal in ISO terms), who is the natural person whom the data is about. Usually, individuals who cannot be identified or distinguished from others are not considered as data subjects.
- A Data Controller (or PII Controller) which determines the purposes and means for processing the data. A data controller must be a natural or legal person, so it can be held accountable if required, and must ensure that any processing of personal data for which it is responsible complies with regulation.
- A Data Processor (or PII Processor), who processes information on behalf of the Data Controller.

In addition, the literature identifies different privacy stakeholders i.e. natural or legal person, public authority, agency or any other body that can affect, be affected by, or perceive themselves to be affected by a decision or activity related to data processing. For example, the Data Protection Act defines a Data Protection Authority as the independent legal authority that administers privacy rules within a country.

We think these roles may well be the starting point to allocate privacy requirements, although it would be necessary to ensure first that the set of roles is complete.

C. Structuring Privacy Requirements into Layers

The idea of refining principles into lower-level, more detailed goals is not new to the privacy realm. As early as in 2001, Langheinrich [16] proposed a decomposition of PhD principles into so-called guidelines, although they were only applied to a specific domain (ubiquitous systems). These days, there have been different attempts to refine privacy principles into more detailed points, which demonstrate our proposal fits within the current state of the technique, where gaps need to be nonetheless covered.

The OASIS Privacy by Design Documentation for Software Engineers Technical Committee (PbD-SE) [17] is working on “the specification of a methodology to help engineers model and document Privacy by Design (PhD) requirements, translate the principles to conformance requirements within software engineering tasks, and produce artifacts as evidence of PhD-principle compliance”. PbD-SE, chaired by Cavoukian herself, splits PhD principles into sub-principles, and later into detailed conformance requirements. We agree with their consideration that better documented privacy requirements would enable software engineers in issuing proofs of normative judgments, privacy claims, and regulatory compliance, and their decomposition into a second layer of sub-principles could be quite useful for our goals. However, their lower layers are not aligned with our proposal, as PbD-SE conformance requirements focus quite much on the documentation process, instead of the final product requirements themselves.

Cavoukian’s Operationalizing Privacy by Design [6] refines each of the privacy principles into a set of actions and responsibilities, plus it compiles other documents published by her team as well, which contains detailed guidance. These documents are sometimes addressed to concrete roles (e.g. boards of Directors); some others deal with a focused goal or task (e.g. Governance, Risk Assessment); yet others focus on specific application domains (e.g. Health Care, Smart Grid) and technologies (e.g. Mobile Services, Wireless Communications, Big Data, Biometrics, RFID, Federated domains); and finally, some merely reflect best practice cases. As we see, they do not follow a systematic decomposition approach, but they mix up different levels. A clearer, systematic structuring would make much easier to trace requirements, decomposed through the different layers, and evaluate the completeness of a list of requirements.

Different authors have defined hierarchies of privacy patterns that help move from legal principles to technical designs. High-level architectural patterns (e.g. minimization, enforcement) include Hoepman’s strategies [18] and Kung’s tactics [19]. In both cases, each architectural pattern is then refined into low-level design patterns i.e. a reusable solution to a recurrent privacy problem. These patterns describe the problem at stake, the context in which they can be applied, and the implementation of the solution, which might be a useful technique to refine more abstract requirements.

The architectural patterns can be mapped onto one or more legal principles, some of which can only be fully covered with additional procedural and organizational means. The low-level design patterns are useful to streamline the design process, since they can be reused by developers (we will next revisit those, when dealing with techniques). However, they are not enough on their own, as they jump into the design without refining the requirements themselves.

On the other hand, our proposal finds a paragon in existent non-functional requirement modeling frameworks. For instance, Yu and Cysneiros [20], employed i* to operationalize privacy requirements, moving from abstract requirements (softgoals), to concrete requirements (goals) and design solutions which operationalize these requirements (tasks). This approach also allows conjugating different definitions of privacy, from the perspectives of different agents playing different roles, so it seems quite promising to inspire the organization into requirement layers we propose.

Fig. 2 summarizes the current landscape compared to our proposal: Although some existent proposals address the overall concept of decomposition, they usually mix up organizational and technical requirements altogether, without providing low-level, specific technical requirements. Considered individually, they leave gaps in the decomposition process. However, a comprehensive set of requirements could be derived from all these and other sources, and then be structured according to our proposal. As a matter of example, under Cavoukian’s principle of “Full Lifecycle Protection”, a guideline for data processors would be to “Employ strong encryption by default along the different processes”, a corresponding success criterion would read “encryption keys must be of sufficient length to resist
breaking attempts and remain protected from access by unauthorized individuals," a general technique would match Hoepman’s encryption design pattern, and a technology specific technique could be a technical mean to implement this pattern e.g. using Advanced Encryption Standard to encrypt all personal data. Of course, some purely legal requirements would not map to this decomposition, as they do not introduce technical requirements (e.g. registration of personal information repositories at data protection authorities).

D. Technique Sources

Privacy Enhancing Technology (PET) is a widespread term to describe a set of measures, products, or services useful to improve one or several aspects of privacy protection, usually in specific contexts. This term groups disparate solutions, which are described at different levels of abstraction and focus on meeting different privacy principles. Thus, they do not help to refine the requirements but to meet them once they are stated.

The lack of a systematic methodology to understand which PETs better fits with a given privacy requirement and context makes selecting one among them a difficult task. Hoepman [18] has tried to generalize and group PETs in privacy patterns. Likewise, the PriS method introduces process patterns to satisfy privacy goals that can be tailored according to organizational goals (functional requirements), which have been successively decomposed into a hierarchy of subgoals [21]. However, there is no current agreement on the way to describe privacy patterns, the level of abstraction they should provide, or how to choose among them. The PRIPARE project (http://pripare.eu/) is defining a privacy pattern template which may help homogenize the descriptions, in order to collect and classify them so as to create a privacy pattern repository that will help to choose the right pattern to meet specific privacy requirements, thus contributing to refine privacy requirements.

On a different perspective, the OASIS PMRM, ISO/IEC 29101 and the NIST Privacy Controls define a set of privacy services that every system should include as part of their privacy architecture. For example, the OASIS PMRM identifies 8 privacy services logically grouped into Core Policy, Privacy Assurance, and Presentation and Lifecycle services. The PMRM methodology proposes conducting a detailed use case analysis to identify the privacy requirements, and then use the privacy services to meet them. However, there is neither guidance on the requirements to be met (but merely high level privacy principles), nor how the services must be used to fulfill them.

E. Levels of Conformance

The concept of different levels of privacy protection is already adopted by different legislations, which usually attach it to different degrees of sensitivity of personal information: different types of personal information require different measures (e.g. UK Data Protection Act, Australian Data Privacy Act as amended, Spanish regulations in force according to the Personal Data Protection Law, US Department of Homeland Security Handbook for Safeguarding Sensitive Personally Identifiable Information). Following our proposal, these different measures would be mapped to different subsets of success criteria, each assigned to a respective level of conformance.

In addition, a clear definition of levels of conformance would enrich two trends in the state of practice in privacy assessment: privacy seals and privacy level agreements.

A privacy seal is the outcome of a certification procedure regarding the privacy compliance of processes, technologies, products and services. It usually includes the evaluation of legal and, very few times, technical standards. As conclude by Rodrigues et al. [22], who comprehensively inventoried and analyzed 25 privacy and related certification schemes in Europe and at the international level, the privacy seals landscape is heterogeneous since there is a large degree of variation around the core functional models of seals schemes, the features they check, the level of guarantees to individuals, etc. Clearly defined levels of conformance could help this plethora of, sometimes incompatible, privacy seals converge to a compatible set, albeit possibly including different levels of certifications (different seal requirements).

A privacy level agreement (PLA) [23] addresses the level of privacy protection a service provider commits to undertake and maintain - they represent the privacy-specific part of a Service Level Agreement (SLA). A PLA answers a lengthy, standardized set of questions, detailing policies and measures the service provider declares to put in practice. Thus, it represents an effective way for service providers to communicate to their customers the level of data protection they may expect. However, the large amount of details it provides may lead to a great variance which may eventually disorient customers. If these PLAs were supplemented by an assignment to one of a set of levels, they would be perceived much clearer to customers.

IV. DISCUSSION AND CONCLUSIONS

We have introduced five conditions that should be met in order to map the requirements framework used in accessibility onto the privacy realm. In most cases, these could rely on the state of the art, by taking further steps to restructure the existing work or filling the currently uncovered gaps. However, there is a key point that goes beyond the technical aspects: the need for a common standardization forum, agreed among all the stakeholders, which could issue standards accepted by all of them. We think OASIS is in good position to assume that role, yet further convergence is still needed with the community outside, and among their own working groups.
Adopting such a framework would be of much use, as the success of WAI for accessibility has already proven. Certification mechanisms would flourish based on that common framework. A consensus on requirements could even impact regulations worldwide, by homogenizing them. Besides, a proper definition of requirements for a limited domain could eventually be extended to any other scenario.

Nonetheless, nobody says this is a seamless task. A major, philosophical difference exists between accessibility and privacy: the former is just a matter of Human-Computer Interaction, while the latter is a Socio-Technical Systems issue [24], where different social entities interact with each other and with technical subsystems, and these interactions cannot be reduced to mere computing system issues. This has implications for our framework: requirements for each role would not probably be so clearly separated, and assessment from a mere user perspective would be difficult -there is not a single input-output user port anymore, but several entities interacting with the system in the back-office. Another drawback is the different meaning attached to privacy around the world. While we recognize that an agreement on privacy basics within the community is needed, we also consider how the original differences between different approaches to accessibility (e.g. barrier removal and closed assistive technologies versus universal design) were reconciled while trying reaching a common requirement corpus -a successful experience that we precisely aim at translating.

In any case, the advantages seem to compensate for the potential drawbacks, and we consider this idea could deserve further study, with any needed adaptations to the privacy realm, to ease the tasks in privacy requirements engineering.

ACKNOWLEDGMENT

This research has been partially supported by the PRIPARE project funded by the European Union’s Seventh Framework Programme under grant agreement number ICT-610613.

REFERENCES