Towards a Roadmap for Privacy Technologies and the General Data Protection Regulation: A Transatlantic Initiative

Stefan Schiffner[1](✉), Bettina Berendt[2], Triin Siil[3], Martin Degeling[4], Robert Riemann[5], Florian Schaub[6], Kim Wuyts[2], Massimo Attoresi[5], Seda Gürses[2], Achim Klabunde[5], Jules Polonetsky[7], Norman Sadeh[8], and Gabriela Zanfir-Fortuna[7]

[1] University of Luxembourg, Esch-sur-Alzette, Luxembourg
Stefan.Schiffner@uni.lu
[2] KU Leuven, Leuven, Belgium
{Bettina.Berendt,Kim.Wuyts,Seda}@kuleuven.be
[3] Cybernetica, Tallinn, Estonia
triin.siil@cyber.ee
[4] Ruhr-Universität Bochum, Bochum, Germany
Martin.Degeling@ruhr-uni-bochum.de
[5] EDPS, Brussels, Belgium
{Robert.Riemann,Massimo.Attoresi,
Achim.Klabunde}@EDPS.europa.eu
[6] University of Michigan, Ann Arbor, USA
fschaub@umich.edu
[7] Future of Privacy Forum, Washington, USA
{julespol,gzanfir-fortuna}@fpf.org
[8] Carnegie Mellon University, Pittsburgh, USA
sadeh@cs.cmu.edu

Abstract. The EU's General Data Protection Regulation is poised to present major challenges in bridging the gap between law and technology. This paper reports on a workshop on the deployment, content and design of the GDPR that brought together academics, practitioners, civil-society actors, and regulators from the EU and the US. Discussions aimed at advancing current knowledge on the use of abstract legal terms in the context of applied technologies together with best practices following state of the art technologies. Five themes were discussed: state of the art, consent, de-identification, transparency, and development and deployment practices. Four traversal conflicts were identified, and research recommendations were outlined to reconcile these conflicts.

1 Introduction

The European Union's General Data Protection Regulation (GDPR), effective as of May 2018, constitutes, intentionally, a big challenge: how to bridge the gap between the fundamental rights it aims to protect including the legal reasoning and instruments proposed to effect such protection on the one hand, and the technologies that threaten

M. Medina et al. (Eds.): APF 2018, LNCS 11079, pp. 24–42, 2018.
https://doi.org/10.1007/978-3-030-02547-2_2

and/or serve to protect these rights including the technological-economic reasoning and practices that give rise to these technologies on the other hand. The GDPR also presents further challenges, such as how to make it work in a worldwide context in which those who control and process personal data are often from legal and social cultures different from those of the EU, with the exact nature of the differences itself a topic of continuing debates. The present paper arose from a workshop under the heading of the second challenge: "Privacy Engineering and the GDPR: A Transatlantic Initiative." For the "EU" and "US" backgrounds that participants represented during the workshop, we found much more unifying than dividing legal and social concerns. The workshop thus constituted the beginning of a transatlantic initiative for drawing a roadmap to address the challenge.[1]

To address the complexity of the questions, we formulated sub-problems, with a view to letting these answer the challenge in conjunction. Five key themes were identified as crucial in the roadmapping process: (1) what is the state of the art in technology; (2) how can consent be meaningful; (3) is de-identification a usable tool; (4) how can processing be transparent and interpretable; (5) what are further challenges given current development and deployment practices?

The challenges concern both the *what* and the *how* of design: the question of what technologies are regulated and/or required by the law, and the question of how the legal requirements can be mapped to software design processes.

Regarding the *what*, GDPR Article 2(1) clearly points out that it applies to all operations (in GDPR terminology, 'processing') performed on any information relating to an identified or identifiable natural person (in GDPR terminology, 'personal data'), whether carried out by automated means or manually. As further explained in GDPR Recital 15, this reflects the principle of technology neutrality – the protection of natural persons should not depend on the techniques used. The natural persons who are identified or identifiable by means of personal data are referred to as 'data subjects' in GDPR terminology. GDPR further conceptualizes a 'controller' as a person who decides why and how personal data is processed and a 'processor' as someone who processes personal data on behalf of a controller.

The GDPR addresses technologies for data processing in two main contexts:

1. GDPR Article 24(1) requires controllers to implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing of personal data is performed in accordance with the GDPR. This requirement is further specified in other norms of the GDPR: (a) data protection by design – Article 25(1) requires appropriate technical and organisational measures, such as pseudonymization, designed to effect data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing. (b) data protection by default – Article 25(2) requires appropriate technical and organisational measures for ensuring that, by default, only personal data that are necessary for each specific purpose of the processing are processed. (c) security of processing – Article 32(1)–(2) require appropriate technical and

[1] Cf. also the US National Privacy Research Strategy, https://www.nitrd.gov/cybersecurity/nationalprivacyresearchstrategy.aspx.

organisational measures to ensure a level of security appropriate to the risks presented by processing (accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data transmitted, stored or otherwise processed).

2. By applying certain technical measures, the data processing may be partly or entirely out of the scope of the GDPR. With reference to GDPR Article 11, a controller may process personal data under relaxed terms and conditions if it took measures to sanitize data in such a way that data subjects cannot be identified. Furthermore, the GDPR does not apply to anonymous information (Recital 26). In both cases, the GDPR does not specify which technologies should be considered as appropriate. Nevertheless, the GDPR requires controllers and processors to give due regard to the state of the art when choosing the technologies (GDPR Articles 25(1) and 32(1)).

For both requirements, uncertainty persists over the available technologies, the burden their adoption adds for controller and processor, their actual functionality (protection goals), and their relation to the legal requirements.

Regarding the *how*, mappings from legal requirements to available technologies have been proposed by several authors. Particularly relevant for the present purposes are those that map from GPDR principles, first to the software design process (design patterns) and from patterns to available technologies [1]. However, the spectrum of available technologies is heterogeneous (in readiness and provided functionality) and volatile (frequent innovation and obsolescence of technology). In addition, the mapping sequence stresses a top-down or even waterfall-model design, which neglects actual practices as well as state-of-the-art insights about agile design methods. While there are initiatives to establish and maintain a repository of technologies ([2] and IPEN[2]), this bottom-up view and the needs of an agile design process remain under-researched.

The contribution of this paper is a roadmap outlined via five key themes and four transversal conflict areas. The procedure that led to these themes, the workshop and its discussion groups are described in Sect. 2, and the theme-specific results in Sect. 3. In Sect. 4, we present four transversal conflicts that we have identified through a synopsis of the themes. From these conflicts, in Sect. 5 we then derive conclusions for future research.

2 Background: Transatlantic Initiative and Workshop

Motivated by the GDPR going into full effect in May 2018, the Internet Privacy Engineering Network (IPEN), the Future of Privacy Forum, the KU Leuven (University of Leuven) Computer Science Department/DTAI group, and Carnegie Mellon University's Privacy Engineering Program decided to host a joint workshop in Leuven, Belgium, as a transatlantic initiative. With this November 2017 workshop, we aimed to determine the relevant state of the art in privacy engineering with a focus on those areas where the "art" needs to be developed further.

[2] www.engineeringprivacy.eu.

The goal of this initiative was to identify open research and development tasks that are needed to make the full achievement of the GDPR's ambitions possible. Within this thematic scope, we wanted to focus the workshop on those areas most relevant to our envisaged audience. We spread an informal survey in our networks, asking people "for a quick shortlist of the most pressing issues (a) that [they] have encountered in the preparation for adapting [their] data processing to the GPDR in their organisation, (b) that [they] perceive as a gap, or as getting too little attention, in research, whether in [their] particular area or in any other, and that [they] think privacy engineers should focus their attention on."

We received a wide range of insightful input from more than 40 people from academia, industry, civil society and regulators, mostly but not exclusively from computational and legal backgrounds. We grouped the answers into categories and selected those that had garnered most interest. The resulting five themes (see Sect. 3) were described in a call for contributions that was sent out via the same (and additional) channels as the survey, asking potential participants to describe their interests with respect to these themes and to outline how they and the workshop could profit from one another.[3]

This resulted in a group of 105 participants whom we invited to come to the workshop. We were happy to see the intended balances along a number of dimensions, in particular EU/US and academia/industry/civil society/regulators. The program reflected this balance: opening statements by the European Data Protection Supervisor were followed by a keynote presentation and a panel discussion involving researchers, industry and standards bodies representatives. Five breakout groups then worked on the themes and presented their results in the forum.[4]

3 Privacy Engineering and the GDPR: Five Key Themes

In this section we discuss the results of the working groups: (1) what is the state of the art in technology; (2) how can consent be meaningful; (3) is de-identification a usable tool; (4) how can processing be transparent and interpretable; (5) what are the further challenges from development and deployment practices.

3.1 What is the "State of the Art" in Privacy Enhancing Technologies?

The guiding questions of this theme were: How is the state of the art of privacy engineering defined and who defines it? What PET tool boxes can be used for developers, corporate decision makers and supervisory bodies? What data-driven risk assessment frameworks for implementing Privacy by Design in data science and big data analytics already exist? How can these be improved?

The GDPR mandates, in Article 25 on *data protection by design and by default*, controllers of data processing to take into account among others the *state of the art*

[3] https://fpf.org/2017/08/30/privacy-engineering-research-gdpr-trans-atlantic-initiative/.

[4] https://fpf.org/wp-content/uploads/2017/08/TransAtlantic-GDPR-Workshop-Agenda-1.pdf.

when defining means for data processing and during the data processing itself. While the state of the art is also mentioned in Article 32 on *security of processing* and in Recitals 78 and 83, a definition comparable to those in Article 4 for e.g. personal data or processing is not given.

Furthermore, the requirement to employ state of the art technologies to protect personal data is not an absolute requirement. According to Articles 25 and 32, it is to be balanced with the "costs of implementation, the nature, scope, context and purposes of the processing as well as the risks [...] and severity for the rights and freedoms of natural persons" posed by the processing.

Controllers and processors in charge to ensure compliance with the GDPR have to determine for their respective means of data processing which state-of-the-art they have to take into account. This is so far a difficult task, because of the missing definition of the state of the art and the unavailability of guidance and case law on this matter and a lack of experience as the GDPR is still new. With no body by law in charge of establishing the state of the art, data controllers, DPAs, self- and co-regulatory bodies, and EU courts will have to determine the minimum requirements case by case.

Technology, and as such the state of the art, are subject of continuous research by public and private actors and evolve in time. As a result, compliance taking into account the state of the art is a moving target. Emerging new technology may increase the risk of data breaches throughout the life time of a product or service. For instance, the availability of faster and cheaper computing resources may allow attackers to break data encrypted in the past much faster. To ensure a constant low risk level of data breaches, the encryption of already encrypted data must be strengthened over time taking into account the current state of the art.

This consideration leads to another set of questions. How can products and services receive security updates, and for how long must updates be provided? How can the controller be sure to not miss out on relevant developments of the state of the art? Who is liable if the controller or processor discontinue their business activity before the end of the product lifetime?

One can also expect interferences with intellectual property law and competition law. For instance, consider a state-of-the-art privacy engineering tool that is proprietary and only offered by a single vendor to competitors under abusive conditions. This situation can be compared to expensive patented pharmaceuticals providing the only cure for certain diseases. Different though here is that those competitors not adopting the state-of-the-art proprietary or non-proprietary privacy enhancing technologies (PETs) may be sued for unfair competition.

To balance the efforts towards data protection and risks for data breaches, but also for privacy risk assessment, the risk must be measured and quantified in the first place. Workshop participants suggested that standardisation and privacy design patterns may simplify this difficult task, which may eventually even enable automated risk assessment. An automation would also benefit the continuous re-assessment of risks throughout the life time of a product or service. Today, different concepts and methodologies exist that make it possible to break down legal high-level requirements to low-level software requirements to be implemented, for example using PETs from common repositories [2]. Participants expressed the need to further streamline, complete and ease such approaches. Without extensive guidance and ready-to-use building blocks, especially small and medium enterprises with no or small research and development teams may struggle to take into account the state of the art.

3.2 Consent

The guiding questions of this theme were: There are detailed parameters for obtaining valid consent under the GDPR and the future ePrivacy Regulation, creating important challenges for sectors such as advertising technology, mobile apps, connected cars, and smart devices. What can engineering contribute, and what should solutions look like?

A data subject's consent is one of six legal bases for data processing defined in GDPR Article 6. In order to be valid under the GDPR, consent must be specific, informed, freely given and unambiguous. Article 7 requires that if data processing is based on consent, the data controller must be able to demonstrate that consent was obtained and freely given by the data subject, that consent declarations need to be clearly separated from other written declarations such as terms of service and must be presented "in an intelligible and easily accessible form, using clear and plain language." Furthermore, Article 7(3) grants the data subject the right to revoke consent at any time and it should be "as easy to withdraw as to give consent." Article 8 further limits the age of consent to age 16 and above, requiring parental consent for children under 16 for the provision of information society services. Data controllers are charged to "make reasonable efforts to verify [...] that consent is given or authorised by the holder of parental responsibility over the child, taking into consideration available technology."

The GDPR requirements for consent pose crucial challenges for privacy engineering. For one, there is a palpable risk that companies will inundate users with consent requests for all data practices to ensure that GDPR consent requirements are met. However, while it plays an integral role in the GDPR, consent is not the only legal basis for data processing. Companies, engineers, lawyers and regulators need to make deliberate decisions about when consent is required and when other grounds for lawful processing suffice. Of particular relevance are contractual relationships (Article 6(1)(b)) and legitimate interests of the data controller or third parties (Article 6(1)(f)), the latter still granting data subjects the right to object, thus opting out of a certain kind of processing. Thus, explicit consent should be used in cases where it is actually required as a legal basis for processing. There is a need for clear guidance and decision frameworks for helping controllers, which may include technology designers, determine on what legal basis for processing to rely on in what situation.

A second challenge is the provision of consent requests in "intelligible and easily accessible form, using clear and plain language." While usability of privacy notices and controls has been studied extensively (e.g., [3]), few ideal solutions exist. Obtaining truly informed user consent remains a challenge, further spurred by the GDPR's important but extensive transparency requirements (Articles 13 and 14). When providing consent prompts and consent-relevant information to data subjects, the level of granularity and specificity of information provided needs to strike a balance between the requirement for clear and plain language, conciseness and the transparency requirements. It is by now widely recognized that consumers rarely read privacy policies and struggle to understand them [4–7]. While they may not reach the same complexity as privacy policies, there is a risk that consent prompts may be stuffed with information in order to satisfy transparency requirements, resulting in yet more dialogs users ignore, and thereby negating the intent of ensuring that obtained consent is informed and an explicit expression of agreement rather than perfunctory. Instead, consent prompts should contain only information directly relevant for informing the consent decision in the given context with more extensive information being available in additional notice layers [3]. Consent prompts should be designed in a user-centric process involving active and extensive user testing. Writing and designing privacy notices and consent prompts should not be an art form but rather follow an evidence-driven process.

Ideally, well-tested and validated consent prompts and user experiences would be shared as best practice design patterns among the community of legal, technical and regulatory privacy professionals. Standardization of consent language and terminology may also be desirable to ensure that legal concepts are adequately represented in language that is clear and understandable to users and data subjects. This would further allow consistent translation of consent-specific text into different languages.

The notion of consent faces a particular challenge in the context of Internet of Things technology, when devices and technology become integrated with the physical environment. In smart homes or other smart infrastructures, one person might set up a device and consent to data collection and processing, but other individuals might also be subject to data collection in a shared physical environment. Future consent solutions will need to become user-centric rather than being focused on specific devices and technology. Can and should consent proliferate across multiple devices of the same user or across different infrastructures in which the user interacts? Can personal devices act as an agent for managing the user's privacy preferences across contexts and "consent" for the user by translating privacy preferences into responses to consent prompts?

A practical challenge is keeping track of user consent to specific data practices and across different entities. What is an appropriate technical record of consent? How should consent with third parties be handled? For example, when third party widgets are presented on a website, is it sufficient if the first part collects user consent or are separate consent records needed for each party collecting information on the website?

A related challenge is the need for companies to accept and respect when users say no to specific consent requests. Consent needs to be bound to a purpose for processing – but how specific must purposes be to be meaningful both for data subjects and organizations?

Questions regarding consent also extend beyond initial consent requests. Once consent has been given (or not), data subjects need to be given opportunities to access data about them, as well as change prior consent indications in order to exercise their right to object (GDPR Article 21). Such ex post controls require closer attention to ensure that they are usable and enable users to effectively review and change their privacy settings and consent expressions.

Article 21(5) states that "data subjects may exercise [their] right to object by automated means using technical specifications." Prior technical privacy specifications, such as P3P [8] or Do Not Track [9], lacked legal basis as well as industry adoption. Article 21(5) creates a legal basis for using and respecting automated technical specifications, but in order to get companies to adopt a certain specification, they may need assurance that a given standard is considered compliant with Article 21(5). Thus, research and development of privacy agent solutions and technical standards needs to be accompanied by legal assessment and possibly certification to facilitate adoption and support of automated privacy management approaches by industry. Certified GDPR compliance for consent technologies would provide a strong incentive for companies to adopt more innovative and more effective consent approaches. Such certification should also be based on actual user testing to ensure that consent prompts and processes are understandable and usable.

3.3 De-identification

The guiding question of this theme was: How can different levels of de-identification techniques be used or further developed to effectively advance the obligations under the GDPR?

Anonymization and pseudonymization are frequent topics among both academics and practitioners, since the GDPR provides several new provisions relating to these topics compared to the earlier Data Protection Directive (Directive 95/46/EC). The GDPR explicitly introduces the term ‘pseudonymization’ in Article 4 p. 5. If a controller or processor de-identifies personal data, i.e., makes it hard to link such data to an identifiable person, they are permitted to process such “depersonalized” information under relaxed terms according to Article 11 of the GDPR. If they can demonstrate the “depersonalized” information is “anonymous” in terms of Recital 26, the GDPR requirements on data protection do not apply. The Article 29 Working Party (hereinafter “Art. 29 WP”) in [10] declares clearly that “anonymised data do fall out of the scope of data protection legislation.” Hence, de-identification techniques most likely are an important aspect of GDPR compliance. Nevertheless, the GDPR does not provide guidelines how to achieve de-identification. Due to this lack of implementation guidelines, data controllers and processors who choose the opportunity of using de-identification techniques face a risk of future enforcement action. For this section, we outline the obstacles discussed in our breakout session.

A first obstacle are mismatching terms. We often observe, in discussions among legal and technical experts, that certain terms are used very specifically on one side, but understood very broadly on the other side. De-identification is one example: while the terms “anonymous” and “pseudonymization” used in the GDPR seem to be overarching technology neutral terms for lawyers, technologists tend to treat them as subsets of a broader concept of “de-identification.” It was discussed in the group that the GDPR terms “anonymization” and “pseudonymization” do not seem to adequately cover all available de-identification techniques from a technical perspective. PET providers, academics, think tanks and data protection authorities who were represented in the group had very different background knowledge and views on this topic, influenced by different levels of understanding of and access to current state of the art in PETs and privacy engineering skills. Consensus is still a work in progress in this area.

For the rest of this section, we will use the Art. 29 WP definition: “anonymization constitutes a further processing of personal data; as such, it must satisfy the requirement of compatibility by having regard to the legal grounds and circumstances of the further processing." [10] We conclude from this that the terms anonymization and pseudonymization in the GDPR are meant by Art. 29 WP to refer to database sanitization and explicitly not to anonymous communication or other techniques of de-identification (e.g., collecting data in anonymous form, encryption, secret sharing, multi-party computation), even if this may seem counter-intuitive in light of the technology neutrality principle provided for in Recital 15. Despite this, we were able to agree in our discussion group that the terms "anonymous" and "pseudonymization" are used consistently in the GDPR.

However, many participants pointed out that the intended scope of the terms "anonymous" and "pseudonymization" remains unclear. The terms seem to be setting too abstract a goal to be directly implemented in IT practices. Further, it is unclear to which standard the efforts to sanitize data from personal information should be held. In particular, what happens if personal data gets exposed despite sanitization efforts? For practitioners, this uncertainty brings along a risk of legal actions. Hence, there were frequent requests in the discussion group for more guidance on how to achieve compliance with GDPR.

In an attempt to provide such guidance on which tools can be used, Art. 29 WP elaborates on some of the technical means and properties that can be used for achieving anonymization, namely noise addition, permutation, differential privacy, aggregation, k-anonymity, l-diversity and t-closeness. [10] These techniques are evaluated from a qualitative angle. However, from an impact or risk-assessment point of view, this leaves open the question of quantification. How can the level of anonymization be measured, and which level is considered appropriate?

In quantification discussions, it is easy to settle in extreme positions. On one hand, the limitations of current anonymization techniques are often pointed out.

Indeed, it is easy to find proof-of-concept attacks on published databases that were supposed to be anonymised, e.g. [11]. Moreover, when speaking in absolute terms, it is quite likely that there is no method to effectively achieve 100% de-identification of personal data. This could lead to the conclusion that database sanitization is futile.

On the other hand, downplaying the situation by claiming that only a small number of data items have been re-identified in proof-of-concept studies, as e.g. in [12], neglects the nature and motivation of these studies. The aim of studies on de-anonymization is to demonstrate the effectiveness of a certain statistical method, not to mount an actual attack. It needs to be pointed out that an actual attack (1) would often not qualify for a scientific publication and (2) the attacker very likely would rather keep their knowledge private. Hence, deriving attack success rates from published proof of concept attacks is unrealistic.

Either of these extreme positions would lead to less privacy. In the worst case, this leaves no incentive for data controllers and processors to make any efforts in that direction. At the same time, it is clear that even the simple deletion of direct identifiers provides some protection, for example from leaks of small parts of the data or from accidental identification. While the "publish and forget" mentality of the other extreme will not work either, there is a middle ground, as [13] conclude: "the GDPR is compatible with a risk-based approach when contextual controls are combined with sanitization techniques."

3.4 Transparent and Interpretable Processing

The guiding questions of this theme were: How can data mining and machine learning methods be made transparent and interpretable? What exactly should be revealed and how? How can we ensure these methods correspond to GDPR requirements and are understandable to the relevant groups of users?

The requirement that (especially AI-based) decision making or decision support systems provide explanations is as old as expert systems themselves, the value of different types of explanations for different audiences have been studied empirically, e.g. [14], and the call for transparency of algorithms is likewise not new[5]. However, the urgency of these desiderata has increased tremendously with the increase of applications and also in complexity of machine-based or –assisted decisions. The GDPR declares transparency to be a guiding principle of all (personal-data) processing (Article 5(1)(a), [15]), and it requires data controllers to provide specific types of information concerning data held and processing performed, as well as "meaningful information about the logic involved" (Articles 13–15). Further principles, such as accountability (Article 5(2), which requires processes and documentation that ensure and show that and how the data are protected), and rights, such as data portability (Article 20), can enhance transparency and interpretability. Recitals 63 and 71 specify requirements that are related to understandability/interpretability.

There has been intensive discussion in the literature between computer science and law just what it is that the GDPR requires – whether this amounts to a "right to explanation" or something less/else [16, 17], and what the specifics of an "explanation" are [18–20]. Is it general information about how a decision-making system works (the system logic, the data categories), or is it an explanation of individual decisions that such a system makes? Does it include justification of a decision or just the mere facts of the results and its effects? The recent Guideline by the Art. 29 WP [21, p. 26] tends to support the latter perspective as it states that "[the data controller] provides details of the main characteristics considered in reaching the decision, the source of this information and the relevance". But while this might be suitable for conservative credit scoring systems or health related information systems, it seems inapplicable to many big data applications that use a high number of characteristics from various sources and might make not one but many decisions repeatedly. Apart from being difficult to resolve in the individual disciplines, these questions are a clear example of the difficulties of mapping between legal and computational notions.

What does it mean for information to be *meaningful* for achieving transparency and interpretability/understandability? To this question raised by [17], we want to add: Is "more" explanation always "better"? Or is it possible that explanations are vulnerable to the same unreflected big-data assumptions as processing itself? Would people (data subjects as well as other stakeholders such as monitoring agents) just ignore additional inundations of information, as they do with privacy notices [22] and consent prompts (see Sect. 3.2)? Similar arguments are made about breach notices [23], as one indicator of 'the limits of notice and choice' [6]. Could an explanation itself become a leakage

[5] E.g. 30 years of EPIC's work: https://www.epic.org/algorithmic-transparency/.

channel that endangers the protection of personal data? It has been observed in privacy-preserving data mining that for example machine-learned rules, which are good candidates for explanations, can do this, e.g. [24].

In addition, an argument like the following could also be made about explanations: “Notices are always a second-best tool because they only *respond* to breaches, not *prevent* them. Moreover, they shift the burden from the responsible parties to the innocent data subject.” [23, p.1]. Linking interpretability with intervenability [25] can be a partial remedy here, as it may reduce the burden on the individual by limiting the necessity of awareness to cases in which decision systems produce unexpected results. [26] This line of thinking highlights that transparency and interpretability always need to be seen in the context of all principles of data protection.

Another possible obstacle to being meaningful is an over-emphasis on algorithmic or technological aspects. Data protection by design includes the implementation of “appropriate technical and organisational measures” for implementing the GDPR’s principles (Article 25). *How* is the meaningful information actually given to data subjects or other questioning parties? How does this relate to the requirements of accountability? Especially in the computational literature, questions of human-computer interfaces, organisation, and process tend to be disregarded, but we expect this to become one of the key interdisciplinary challenges for effective transparency and interpretability.

Both technical and organisational measures need to support the GDPR requirements to *provide* or *(enable to) obtain* “meaningful information about the logic involved” (Articles 13(1)(f), 14(2)(g), 15(1) (h)). The legal text thus formulates requirements to do something, which however in the computational literature tend to be discussed as non-functional requirements, e.g. “algorithms should be transparent”, “systems should be interpretable”. How can these be turned into requirements that are functional in the sense of software development? What *exactly* does a socio-technical system need to *do*? In this area, we observe a clear conflict between the need for (or the impossibility of) exactly specified requirements: the legal requirement is intentionally underspecified, but the system designer needs exact specifications. In addition, if we consider interpretability as linked to intervenability, what other functional requirements result from this?

Explanations and interpretability also pose challenges for economic interests because what is meaningful to a data subject may be regarded as a trade secret by the data processor. Data protection authorities would then also need to check for what could be termed “explanation fraud” in analogy with audit fraud as in the recent automobile-exhaust scandal: A system could be designed such that, when prodded for an explanation, it generates an answer that satisfies the legal requirements but that does not accurately reflect the normal workings of the system. When machine learning is employed for explanations (e.g., [27]), what roles will adversarial machine learning and secure learning play?

Future research thus needs to answer the following questions:

- What constitutes a “meaningful” or “sufficiently comprehensive explanation” of an automated decisions process?
- What explanations are suitable for which algorithm and which usage context?

- What are user perspectives on this? Can we find levels of explanations that are understandable for different audiences?
- Can modes of intervention be used to prevent the information overload that we see with privacy notices?
- How can we prevent those explanations from leaking either personal information or intellectual property?

3.5 Challenges Arising from Development and Deployment Practice

The guiding questions of this theme were: How can PETs and data protection by design methodologies be integrated into existing software development approaches (especially agile software development)? With software production and use phases collapsing [28], users are integral to experimentation, developers are users themselves, and usability becomes central. Different requirements may be commensurate, complementary, and contradictory. How can we design and evaluate for users and for a democratic society?

The GDPR requires the implementation of Data Protection by design (DP-by-design) and by default. However, both the open-ended scope of the legal requirement, as well as the challenges associated with incorporating PETs and DP-by-design methodologies into existing software development ecosystem requires further attention from researchers and practitioners.

One concern arising from the integration of DP-by-design in existing software development approaches is situated at the inception phase of this integration, namely the translation between legal obligations and technical requirements. Currently there are some *mismatches between legal and technological terminology and conceptual systems*. Key terms may have entirely different meanings in both fields. For instance, the term ‘data owner’, often used by developers, has no legal meaning, and concepts such as ‘data controller’ and ‘data processor’ are in general unknown by software engineers. A privacy engineering vocabulary and ontology that can be used by all parties involved in the software engineering life-cycle (including DPOs, developers, business owners, product owners, etc.) may help address this matter. Terminological and ontolgical efforts such as those in [29–34] have so far had minor impact due to the prevalence of a legal interpretation of the term ‘personal data’ without reference to its derivation from actual concepts present in computational views on data. A common vocabulary would be very helpful in evaluating, for example, third parties before service integration, or developing GDPR certification schemes, but is also likely to fall short of addressing contextual aspects of privacy.

Moreover, there is a *discrepancy between the non-functional requirements that are provided and the necessary requirements that need to be functional*. Legal requirements are often vague (on purpose) and can be hard to translate into technical requirements. High level privacy requirements may not be straightforward to implement if they only express qualities to be achieved. These may hence be hard to implement and validate. Developers can be supported through the integration of privacy engineering principles and processes into existing tools. While efforts have been put into improving security usability and process for developers, the same efforts in privacy are only commencing.

The move to service architectures, data centric software development and increased use of machine learning in services introduces further challenges. While companies' current belief regarding (personal) data is "the more, the merrier", the concept of *'big data' clashes with GDPR's data minimization* principle. Hence, changing the mindset in the industry seems to be a prerequisite for an effective privacy engineering practice. Today, for companies, their (data) assets are the main driver for integrating security into software development. This is, however, where security and privacy differ: while security has typically been applied for protecting company assets, privacy requires protecting users and their interests. A cultural change would have to start with, for example, companies and developers seeing themselves as responsible for protecting the data subject's rights. This may, however, put developers in the odd position of resolving conflicting user and organizational requirements. This may sometimes be resolved using design, e.g., Privacy Enhancing Technologies are often designed to address seemingly conflicting requirements, but not always.

The position of the developer may be strengthened through DP-by-design best practices. Concretely, better integration of the Data Protection Impact Assessments with DP-by-design as well as the Development Life Cycle could help surface conflicts and allow organizations to consider technical and non-technical ways to resolve conflicting requirements. Certification efforts may help create common practices and support companies when they have to integrate third party services. Advancing these projects may, for example, require the development of domain specific standards for DPIAs or better clarification of accountability and documentation requirements when it comes to DP-by-design efforts.

The accelerated iterative approach typical in agile development environments can make DP-by-design more challenging than in a typical waterfall model. In agile development the design of the system is frequently updated and hence requires frequent iterations for privacy and security assessment. Agility, however, requires quick software development sprints, while privacy analysis is typically a slow and time-consuming activity. In addition, technical privacy assessments are based on the architectural description of the system, but in agile development there is, generally speaking, no grand design up front, and little, if any, documentation of the system. It might be possible to assess and integrate privacy for each feature in isolation, but when these features are combined (composed), as well as when services from multiple parties are integrated, there is no guarantee that the service itself, or the entire supply chain that underlies it, fulfills all the privacy requirements. This is especially the case due to modular architectures that are favored in current day software ecosystems. These also raise serious challenges to determining and managing responsibility with respect to privacy. If a company, for example, provides a data service that can be integrated by different third parties, it can be challenging to identify all the privacy risks for all the different uses of the data by these third parties.

Guidance for developers and companies to identify who is legally responsible for the system, as well as studies on how to adapt existing ecosystem to fulfill these requirements will be of great value going forward. For certification efforts, the current ecosystem implies that these should not be one-off efforts but certification processes that integrate frequent updates to software and the service ecosystem into their evaluation.

Legacy systems may also give rise to privacy challenges. It may be difficult to apply privacy 'by design' to existing systems, or achieve sufficient protections through add-on privacy solutions. It may be the case that some systems will have to be re-engineered completely to comply with current laws.

One idea is to adopt tools and techniques from other domains which confront similar challenges such as aviation and medicine [35]. Privacy *can* be formulated as a 'safety' problem. In this respect, tools such as failure model effects analysis [36], data flow modelling [30, 37], checklists [38] and risk management practices [39, 40] directly address many of the issues raised earlier in this paper regarding development processes. Industrial experience has shown this approach to be fruitful [30, 41]; however, it must be noted that some techniques, specifically checklists, come with a significant cultural investment if to be applied correctly and have been criticized in privacy circles as being less than sufficient to provide substantive protections.

The challenges associated with integrating DP-by-design into the development are not the only reason that these principles are often neglected in industry. Many companies still struggle with the identification and documentation of their data processing activities. Clearly, as long as a company does not fully understand what data they process and where it all is stored, applying DP-by-design software development practices will not yet be their priority. For this to change, it is import that integration efforts are intensified and are promoted in industry.

4 Four Transversal Conflicts

We subjected the results described in the previous section to a synoptic analysis and identified four areas of conflicts that re-occurred throughout the themes. We will describe these as transversal conflicts here. In the subsequent Sect. 5, we will use them to identify recommendations for future work. We observe the following conflicts:

1. Hard-to-reconcile and sometimes incompatible terminological and conceptual systems in the legal and technology communities.
2. An unreflected big data ideology that is based on the perception that more data is better even if there is no clear use for the collected data. This is in conflict with data minimization, where data is only collected and stored if needed for a certain functionality.
3. Dysfunctional economics of by-design paradigms if they are implemented as waterfall design process.
4. (From 2 and 3:) A tension between functional vs. non-functional requirements. Design processes, regardless of whether waterfall or agile, are focused on functionality; however, privacy and data protection requirements are often formulated in a non-functional fashion. For accountability (to demonstrate compliance) and enforceability, these properties need to be transformed into functionalities. For example, to ensure confidentiality of communication, encryption needs to be applied.

The protection of personal data and privacy is a particular challenge. It not only describes a "quality" of a system (like all non-functional requirements do) and specifies constraints on the system's functionalities (like many non-functional requirements do), but it is a quality that often derives from "not doing" something (again, data minimization is a good example). However, design(er)s are often focused on making a system "do" something, i.e. on the functional requirements. A focus on "doing" seems necessary to implement accountability, but it may lead to an arms race of technologies and counter-technologies that is counter-productive for the original goal (a phenomenon that predates the GDPR, cf. [42] on similar spirals in risk management).

5 Conclusions: From Conflicts to Roadmap

As the individual themes' discussions show, the conflicts sketched in the previous section re-occurred throughout, although not always in the same intensity. It is probably too early to extract solutions for these conflicts; instead, they offer a productive basis for future work.

Conflict 1: Mismatching Terminology. If left untackled, this conflict will lead to inconsistent interpretations of the legal text, which will have a long-term impact on the technological solutions and the debate on what *state of the art* and *appropriate* mean here.

Terminology divergences are unavoidable because they reflect the different needs of the communities, and these differences cannot easily be resolved by an ontology. For example, on one hand, compliance cannot be defined as a by-design notion, since legal terms are always defined in their context. On the other hand, the designer needs to take implementation decisions for which they need to guarantee legality. In many fields, checklists and standards are established to resolve this conflict.

While standardization plays an important role also in IT security, IT protection measures appear to experience disruptive changes (black swan events) in their effectiveness more often than other technologies. As an example, consider the almost instant scalability of attacks: once a protection measure has been broken, the breaking can usually be replicated rather easily. This is intrinsically different from physical safety measures, where even if a weakness is discovered, a second break-in can be as difficult as the first attempt. This observation leads to the need to keep continuously updated and maintained descriptions of the state of the art.

Lastly for the terminology conflict there is also good news: the often troubled cultural conflict between US and EU may be much smaller than claimed. The need for privacy protection is the same, as is the hope for business opportunities; the conflict is mostly in legal traditions how to reach the same or similar goals.

Conflict 2: Compulsive Data Hoarding. While creativity and market needs might inhibit the application of by-design paradigms, compulsive data hoarding without functionality in mind is more ideology-driven. A wider public debate on the cost for society of this data hoarding is needed. This includes a cost-benefit analysis of knowledge and ignorance. This is generally not new as it is often done in medicine, when the decision needs to be taken whether a medical test is needed or not.

Conflicts 3 and 4: Dysfunctional Economics of by-Design Paradigms and (Non)-Functional Requirements. Given that it is very hard, if not impossible, to "bolt on" data protection functionalities post-production, by-design paradigms seem to be obvious and inevitable. However, their adoption falls beyond expectation. We observe that the waterfall-like development cycle, which is assumed for current PbD interpretations, ignores the creative process that is needed to develop new functionalities and matching business models. Prototypes to test the potential markets for a new functionality need to be developed fast and in cost-efficient ways. This market need was answered by the adoption of agile design methods, driven by use cases and functionalities. However, privacy and security properties need to be turned into functionalities.

Here, PETs come into the play. For example, to implement the requirement that personal data must not be accessible for unauthorized parties, data needs to be encrypted and users need to be authenticated for access.

Having said that, even formulating privacy properties into functional requirements might fall short. These functionalities also need to be considered in the frequent test cycles during development in a meaningful way. For the truly functional part of the security user story (e.g. an encrypt function), this can easily be expressed as a feature, and there will be someone who requires and oversees this in the development team. For the intrinsically non-functional parts such as "every data item needs to be minimised in correspondence with its purpose", such a continuously involved stakeholder is often missing, since it is mostly an end-user requirement or even more strongly for the *non*-users that are affected by (the lack of) privacy solutions.

The law now shifts this protection interest from the end-user domain into the domain of the data controller, by creating liabilities and thus financial risk. This is the point where legal compliance comes into play and where this conflict feeds back into conflicts 1 and 2: There is a need to translate legal abstract terminology into positive and implementable software requirements.

6 A Summary Roadmap

With the GDPR now coming into force, this paper offers initial insights on development practices and organisational measures, including product lifespan until its organized obsolescence as well as the handling of data from its collection and processing, and onto final deletion. We outline the following recommendations and avenues for further research based on the four conflicts identified in this paper. Disparities in terminology is a persistent issue, especially in light of both domains evolving independently: on the one hand, the law will remain relatively unchanged, its interpretation will change by case law and legal praxis and, as a result, so will the concepts behind the words. On the other hand, IT products and the technology on which they are based will change more rapidly and so will the language used to speak about this technology. Much needed now is a platform able to continuously map current interpretations of the law at state of the art level. Regarding compulsive data hoarding, it is anticipated that this issue will naturally fade away notwithstanding its many promises in the field. Researchers can support this process by demonstrating the poor utility and high costs of such unstructured and purposeless data collections. Finally, further research is needed in the field of dysfunctional by-design economics and the handling of non-functional requirements. Greater focus should be placed on truly interdisciplinary research that is directly communicating with the work on matching terminology.

Acknowledgments. This work was partially funded by the European Union's Horizon 2020 project grant no. 740829 and 778615, the Luxembourg National Research Fund project PETIT grant agreement no. 10486741, the National Science Foundation grant agreements CNS-1330596 and SBE-1513957, under the Brandeis privacy initiative DARPA, AFRL grant agreement no. FA8750-15-2-0277, the Research Foundation Flanders, the Research Fund KU Leuven, and the KUL-PRiSE research project.

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official positions, policies or endorsements, either expressed or implied, of the institutions they are affiliated with, the EDPS, the National Science Foundation, DARPA, the Air Force Research Laboratory or the US Government.

We thank Ian Oliver and Jef Ausloos and the APF reviewers for their valuable input and comments, and all workshop participants for their contributions and stimulating discussions.

References

1. Hoepman, J.-H.: Privacy design strategies. In: Cuppens-Boulahia, N., Cuppens, F., Jajodia, S., Abou El Kalam, A., Sans, T. (eds.) SEC 2014. IAICT, vol. 428, pp. 446–459. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-55415-5_38
2. ENISA: Privacy Enhancing Technologies: Evolution and State of the Art A Community Approach to PETs Maturity Assessment (2016). https://www.enisa.europa.eu/publications/pets-evolution-and-state-of-the-art
3. Schaub, F., Balebako, R., Durity, A.L., Cranor, L.F.: A design space for effective privacy notices. In: Eleventh Symposium on Usable Privacy and Security (SOUPS 2015), Ottawa, pp. 1–17. USENIX Association (2015)
4. President's Council of Advisors on Science and Technology: Big data and privacy: a technological perspective. Report to the U.S. President, Executive Office of the President, May 2014
5. Cranor, L.F.: Necessary but not sufficient: standard mechanisms for privacy notice and choice. J. Telecommun. High Technol. Law **10**, 273 (2012)
6. Cate, F.H.: The limits of notice and choice. IEEE Secur. Priv. **8**(2), 59–62 (2010)
7. Schaub, F., Balebako, R., Cranor, L.F.: Designing effective privacy notices and controls. IEEE Internet Comput. **21**(3), 70–77 (2017)
8. Wenning, R., et al.: The platform for privacy preferences 1.1 (P3P 1.1) specification (2006). https://www.w3.org/TR/2018/NOTE-P3P11-20180830/
9. Fielding, R.T., Singer, D.: Tracking preference expression (DNT) W3C candidate recommendation (2017). https://www.w3.org/TR/2017/CR-tracking-dnt-20171019/
10. Article 29 Working Party. Opinion 05/2014 on anonymisation techniques (2014). WP216. http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf
11. Narayanan, A., Shmatikov, V.: Robust de-anonymization of large sparse datasets. In: 2008 IEEE Symposium on Security and Privacy, SP 2008 (2008)
12. Cavoukian, A., Castro, D.: Big data and innovation, setting the record straight: de-identification does work. In: Information and Privacy Commissioner, p. 18 (2014)

13. Hu, R., Stalla-Bourdillon, S., Yang, M., Schiavo, V., Sassone, V.: Bridging policy, regulation and practice? A techno-legal analysis of three types of data in the GDPR. In: Data Protection and Privacy: The Age of Intelligent Machines, p. 39 (2017)
14. Ye, L.R.: The value of explanation in expert systems for auditing: an experimental investigation. Expert Syst. Appl. **9**(4), 543–556 (1995)
15. Article 29 Working Party. Guidelines on transparency under regulation 2016/679 (2016). 17/EN WP260. http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id = 615250
16. Wachter, S., Mittelstadt, B., Floridi, L.: Why a right to explanation of automated decision-making does not exist in the general data protection regulation. Int. Data Priv. Law **7**, 76–99 (2017)
17. Selbst, A.D., Powles, J.: Meaningful information and the right to explanation. Int. Data Priv. Law **7**(4), 233–242 (2017)
18. Biran, O., Cotton, C.: Explanation and justification in machine learning: a survey. In: IJCAI-17 Workshop on Explainable AI (XAI) Proceedings, pp. 8–13 (2017). http://www.intelligentrobots.org/files/IJCAI2017/IJCAI-17_XAI_WS_Proceedings.pdf#page=8
19. Lipton, Z.C.: The mythos of model interpretability. In: ICML 2016 Workshop on Human Interpretability in Machine Learning (WHI 2016) (2016). http://zacklipton.com/media/papers/mythos_model_interpretability_lipton2016.pdf
20. Edwards, L., Veale, M.: Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for. Duke Law Technol. Rev. **16**, 18 (2017)
21. Article 29 Working Party. Guidelines on automated individual decision-making and profiling for the purposes of regulation 2016/679 (2018). 17/EN WP251rev.01. http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053
22. Obar, J.A., Oeldorf-Hirsch, A., The biggest lie on the internet: ignoring the privacy policies and terms of service policies of social networking services. In: TPRC 44: The 44th Research Conference on Communication, Information and Internet Policy (2016)
23. Cate, F.H.: Information security breaches: looking back & thinking ahead. Technical report Paper 233, Articles by Maurer Faculty (2008). http://www.repository.law.indiana.edu/facpub/233
24. Atzori, M., Bonchi, F., Giannotti, F., Pedreschi, D.: Anonymity preserving pattern discovery. VLDB J. **17**(4), 703–727 (2008)
25. Hansen, M., Jensen, M., Rost, M.: Protection goals for privacy engineering. In: 2015 IEEE Security and Privacy Workshops (SPW), pp. 159–166, May 2015
26. Schmidt , A., Herrmann, T., Degeling, M.: From interaction to intervention: an approach for keeping humans in control in the context of socio-technical systems. In: 4th Workshop on Socio-Technical Perspective in IS development (STPIS 2018) (2018)
27. Ribeiro, M.T., Singh, S., Guestrin, C.: "Why should I trust you?": explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2016, pp. 1135–1144. ACM, New York (2016)
28. Gürses, S., van Hoboken, J.: Privacy after the agile turn. In: Selinger, E., Polonetsky, J., Tene, O. (eds.) The Cambridge Handbook of Consumer Privacy (Cambridge Law Handbooks, pp. 579–601). Cambridge University Press, Cambridge (2018). https://doi.org/10.1017/9781316831960.032
29. Ding, L., Bao, J., Michaelis, J.R., Zhao, J., McGuinness, D.L.: Reflections on provenance ontology encodings. In: McGuinness, D.L., Michaelis, J.R., Moreau, L. (eds.) IPAW 2010. LNCS, vol. 6378, pp. 198–205. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-17819-1_22
30. Oliver, I.: Privacy Engineering: A Data Flow and Ontological Approach. CreateSpace Independent Publishing, July 2014. 978-1497569713

31. Anton, A.I., Earp, J.B.: A requirements taxonomy for reducing web site privacy vulnerabilities. Requirements Eng. **9**(3), 169–185 (2004)
32. Solove, D.J.: A taxonomy of privacy. Univ. Pennsylvania Law Rev. **154**(3), 477 (2006). GWU Law School Public Law Research Paper No. 129
33. Solove, D.J.: Conceptualizing privacy. Calif. Law Rev. **90**(4), 1087–1155 (2002)
34. Kost, M., Freytag, J.C., Kargl, F., Kung, A.: Privacy verification using ontologies. In: ARES, pp. 627–632. IEEE (2011)
35. Kern, T.: Flight Discipline. McGraw-Hill Education, New York (1998)
36. Card, A.J., Ward, J.R., Clarkson, P.J.: Beyond FMEA: the structured what-if technique (SWIFT). J. Healthc. Risk Manag. **31**, 23–29 (2012)
37. Scandariato, R., Wuyts, K., Joosen, W.: A descriptive study of Microsoft's threat modeling technique. Requirements Eng. **20**(2), 163–180 (2015)
38. Gawande, A.: The Checklist Manifesto. Profile Books (2011)
39. Reason, J.T.: Managing the Risks of Organizational Accidents. Ashgate, Farnham (1997)
40. Pfleeger, S.L.: Risky business: what we have yet to learn about risk management. J. Syst. Softw. **53**(3), 265–273 (2000)
41. Oliver, I.: Experiences in the development and usage of a privacy requirements framework. In: 24th IEEE International Requirements Engineering Conference, RE 2016, Beijing, China, 12–16 September 2016, pp. 293–302. IEEE Computer Society (2016)
42. Power, M.: The risk management of everything. J. Risk Finance **5**, 58–65 (2004)