Sarah Spiekermann-Hoff

The Challenges of Privacy by Design

Article (Accepted for Publication)

Sarah Spiekermann*

The Challenges of Privacy-by-Design

Heralded by regulators, Privacy by Design (PbD) holds the promise to solve the digital world's privacy problems. But immense challenges lie ahead: Real management commitment and concrete step-by-step methods to integrate privacy into systems are just the tip of the iceberg.

Privacy maintenance and control is a social value deeply embedded in our societies. A global survey found that 88% of people are worried about who has access to their data; over 80% expect governments to regulate privacy and impose penalties on companies that don't use data responsibly. But privacy regulation is not easy. The Internet's current economics as well as national security management benefit from the collection and use of rich user profiles. Technology constantly changes. And data is like water: it flows and ripples in a way that is hard to predict. As a result, even a well-conceived, general and sustainable privacy regulation, such as the European Data Protection Directive 95/46/EC, struggles to ensure its effectiveness. Companies regularly test legal boundaries and many risk sanctions for privacy breach to avoid constraining their business.

Against this background, the European Commission and other regulatory bodies are looking for a more effective, system- and context-specific balance between citizens' privacy rights and the data needs of companies and governments. The apparent solution, heralded in many legal documents, but barely specified, is "Privacy by Design". At first sight, the powerful term seems to suggest that we simply need to take a few 'PETs' (Privacy Enhancing Technologies) and add a good dose of security, thereby creating a fault-proof systems' landscape for the future. But the reality is much more challenging. According to Ann Cavoukian the Ontario information and privacy commissioner who first coined the term, PbD, stands for a ***pro-active*** integration of technical privacy principles in a system's design (such as privacy default settings or end-to-end security of personal data) *and* the recognition of privacy in a company's risk management processes.[1] PbD can thus be defined as **a pro-active engineering *and* management approach that is committed to selectively and sustainably minimize information systems' privacy risks through technical *and* governance controls.**

However, a core challenge for PbD is to get organizations' management involved in the privacy strategy. Management's active involvement in the corporate privacy

[1] Cavoukian, Anne, 20110 „Privacy by Design Curriculum", URL: http://privacybydesign.ca/publications/

strategy is key because personal data is *the* asset at the heart of many companies' business models today. High privacy standards can restrict the collection and use of data for further analysis, limit strategic options and impact a firm's bottom line. Consider advertising revenues boosted by behavioral targeting practices and peoples' presence on social networking sites: Without personal data, such services are unthinkable. PbD proponents hardly embrace this economic reality in their reasoning. Instead, they take a threat perspective arguing that low privacy standards, can provoke media backlash and lead to costly legal trials around privacy breaches. And indeed, distrust caused by privacy breaches is probably the only real blemish on the image of technology companies such as Google or Facebook. Brands are a precious company asset, the most difficult to build and the most costly to maintain. Hence, brand managers should be keen to avoid privacy risks. Equally, recent data breach scandals have forced CEOs to quit. Despite these developments, many managers still do not understand that a sustainable strategy for one of their company's core assets – personal data – requires them to actively manage this asset. *Managing* personal data means to optimize its strategic use, its quality and long-term availability. Unfortunately, few of today's managers want to take on this new challenge. Instead, they 'milk' what they can from the information bits that they get and leave the privacy issue as a nuisance that is better left to be fixed by their lawyers.

But even if managers took up the privacy challenge and incorporated the active governance of personal data into their companies' strategic asset management, they would not be able to determine the right strategy without their IT departments: PbD requires the guts and ingenuity of engineers. As the term implies, the *design* of systems needs to be altered or focused to technically embrace the protection of peoples' data. Consequently, privacy must be on engineers' requirements radar from the start of a new IT project. It needs to enter the system development life cycle at such an early point that architectural decisions around data processing, transfer and storage can still be made. Managers and engineers (as well as other potential stakeholders) need to assess the privacy risks that they are willing to take and jointly decide on technical and governance controls for those risks that they are not willing to bear.

Even when both managers and engineers are committed to PbD, more challenges must be overcome:

- Privacy is a fuzzy concept and is thus difficult to protect. We need to come to terms on what it *is* we want to protect. Moreover, conceptually and methodologically, privacy is often confounded with security. We need to start distinguishing security from privacy to know what to address with what means.
- No agreed-upon methodology supports the systematic engineering of privacy into systems. System development life cycles rarely leave room for privacy considerations.
- Little knowledge exists about the tangible and intangible benefits and risks associated with companies' privacy practices.

How can these challenges be overcome? A Privacy Impact Assessment (PIA) Framework recently created for RFID technology[2] has been called a 'landmark for PbD' because it offers some answers: The PIA Framework suggests to use concrete privacy protection goals and describes a method to reach them. Pragmatically, it recommends that organizations use the specific legislative privacy principles of their region or sector or the OECD Privacy guidelines *as a starting point* to determine privacy protection goals. In Europe, for example, the European Data Protection Directive 95/46/EC defines the following privacy goals:

- Safeguarding personal data quality through data avoidance, purpose-specific processing and transparency vis-à-vis data subjects
- Ensuring the legitimacy of personal and sensitive data processing
- Complying with data subjects' right to be informed, to object to the processing of their data, and to access, correct and erase personal data.
- Ensuring confidentiality and security of personal data

Security and privacy in this view are clearly distinguished. Security means that the confidentiality, integrity and availability of personal data are ensured. From a data protection perspective security is *one* of several means to ensure privacy. A good PbD is unthinkable without a good Security by Design plan. The two approaches are in a "positive sum" relationship. That said, privacy is about the *scarcity* of personal data creation and the maximization of *individuals' control* over their personal data. As a result, some worry that PbD could undermine law enforcement techniques that use criminals' data traces to find and convict them. More research and international agreement in areas such as anonymity revocation are certainly needed to show that this need not be the case even if we have privacy friendly systems.

After privacy goals are clearly defined, we must identify how to reach them. The PIA Framework[3] mentioned above is built on the assumption that a PbD methodology could largely resemble security risk assessment processes such as NIST or ISO/IEC 27005. These risk assessment processes identify potential threats to each protection goal. These threats and their probabilities constitute a respective privacy risk. All threats are then systematically mitigated by technical or governance controls. Where this cannot be done, remaining risks are documented to be addressed later.

As in security engineering PbD controls heavily rely on systems' architectures[3] Privacy scholars still put too much focus on information practices only (such as Website privacy policies) that are used to 'fix' privacy problems by informing users about them. Instead, they should investigate more how to pro-actively build systems in client-centric ways that maximize user control and minimize network or service provider involvement . Where such privacy-friendly architectures aren't feasible (often for business reasons), designers can support PbD by using technically

2 http://ec.europa.eu/information_society/policy/rfid/pia/index_en.htm
3 Spiekermann, S., and Cranor, L.F. "Engineering Privacy," *IEEE Transactions on Software Engineering* (35:1), January/February 2009 2009, pp 67-82.

enforceable default policies (i.e. 'opt-out' settings) or data scarcity policies (i.e. erasure or granularity policies), data portability and user access- and delete rights. Where such technical defaults aren't feasible, concise, accurate and easy to understand notices of data-handling practices and contact points for user control and redress should come into play.

A challenge, however, is that system development life cycles and organizational engineering processes do not consider such practices. So far, privacy is simply not on an engineer's plate when designing a system. This gap raises many questions: When should privacy requirements first enter the system development life cycle? Who should be responsible? Given that privacy controls impact business goals, who can actually decide on appropriate measures? Must there be ongoing privacy management and practices monitoring? If organizations purchase standard software solutions or outsource operations, pass data to third parties or franchise their brands, who is responsible for customer privacy? How can PbD be audited so that companies can demonstrate their privacy accountability?

For privacy to be embedded in the system development life cycle and hence in organizational processes, companies must be ready to embrace the domain. Here, we return to the starting point. Unfortunately, we still have too little knowledge about the real damage that is being done to brands and a company's reputation when privacy breaches occur. The stock market sees some negligible short-term dips, but people flock to data-intensive services (such as social networks); so far, they don't sanction companies for privacy breaches. So why invest in PbD measures? Will there be any tangible benefits from PbD that justifies the investment? Would people perhaps be willing to pay for advertisement-free, privacy friendly services? Will they incur switching cost and move to competitive services that are more privacy friendly? Would the 83% of US consumers who claim that they would stop doing business with a company that breaches their privacy really do so? I believe that we need to better understand these dynamics as well as the current changes in the social perception of what we regard as private. But research on the behavioral economics of privacy has clearly demonstrated that regardless of what people say, they make irrational privacy decisions and systematically underestimate long-term privacy risks. And this is not only the case for privacy-seeking individuals, but also for managers who are making PbD decisions for their companies. Therefore, I believe that PIAs should be seriously considered in legislation as well as self-regulation. They must however be credible by providing a precise set of rules against which their effectiveness (and legel certainty) can be verified. They must be auditable! If this is the case, PIAs make privacy risks transparent and help companies to take accountable actions. Making them mandatory through EU data protection regulation or in US privacy sector laws or adding them to international privacy agreements such as the EU-US Safe-Harbor Framework (overseen by the FTC) would force us to embrace process-driven, bottom-up ways to embed privacy into code; probably the only way to really protect one of the core ethical values of our Western liberal democracies and constitutions.