A Design Theory for Transparency of Information Privacy Practices
—
Appendix

Authors: Tobias Dehling, Ali Sunyaev; Institute of Applied Informatics and Formal Description Methods, Department of Economics and Management, Karlsruhe Institute of Technology, Karlsruhe, Germany and KASTEL Security Research Labs, Karlsruhe, Germany; {dehling, sunyaev}@kit.edu

A.1 Example of a thought trial on the usefulness of the privacy-by-design framework to inform the design of useful transparency artifacts

The privacy-by-design (PbD) framework proposes seven foundational principles to translate the fair information practice principles (OECD 1980, US Federal Department of Health Education and Welfare 1973) into technocentric goals (Cavoukian 2009). Initially, the PbD framework appeared helpful to inform the design of IT artifacts useful for establishing transparency of information privacy practices (TIPP), but attempting to use it to inform TIPP theory unveiled that the PbD framework was irrelevant for making transparency artifacts useful with respect to consumers' evolving and context-dependent privacy expectations. First, transparency is one of the seven PbD principles and stated as a goal of PbD to make privacy practices more accountable, open, and compliant (Cavoukian 2009). However, the framework does not consider when this is useful for consumers and appears to require the assumption that consumers are always interested in such information. This points to a second, more fundamental issue: the assumption that consumers' privacy expectations are universal and stable, but this is not the case (Altman 1977, Solove 2002) due to changing social and technological conditions (Mulligan et al. 2016). During design and operation of a specific information system (IS), this can be addressed through extensive user studies and better user experience design to account for differences in contexts and consumers' privacy expectations (Rubinstein and Good 2013), but the assumption that privacy expectations can be captured in a universal way in light of contradicting evidence (Mulligan et al. 2016) made the PbD framework irrelevant to development of TIPP theory.

Consumer surveys show, for instance, that consumers' privacy risk perceptions do not correspond to the binary differentiation between information that is or is not personally identifiable underlying the PbD framework (eg, Milne et al. 2017). Legal scholars criticize that data protection laws do not account for relevant social mechanisms that threaten privacy (Peppet 2011) and make plausible arguments that antitrust laws are a more effective foundation to reduce privacy risks than data protection laws, which underlie the PbD framework, because antitrust laws directly target business motives to offload privacy risks and costs to consumers (Day and Stemler 2019). Moreover, findings in computer science, for example, that anonymity is not only conducive to privacy but can also be abused to make attacks on IS more successful (Ahmad and Clark 2021), further challenge fundamental PbD assumptions, for instance, "that it is possible, and far more desirable, to have both [privacy and security]" (Cavoukian 2009, p. 3). Since the PbD framework limits itself to a perspective that appears to be akin to privacy as control (Westin 1967), the PbD framework should be rather thought of as a framework for privacy risk reduction or data protection law compliance by design and is thus useful to set a normative baseline for the engineering of privacy practices (Spiekermann and Cranor 2009), but it is not really useful for building IS that can meet consumers' evolving and context-dependent privacy expectations (Mulligan et al. 2016).

A.2 Examples of thought trials based on empirical studies

Table 1. Examples of ideas and key implications for TIPP theory development from corresponding empirical studies.

#	Idea	Study	Findings leading to rejection of idea	Key implications for TIPP theory	Reference
1	TIPP can be established with a small, standardized set of information on privacy practices.	Development of a privacy notice content ontology based on a literature review of articles mentioning synonyms for privacy notice and content in title, abstract, or keywords.	63 relevant, unique articles remained after assessment of 441 articles discovered by search string in EBSCO, ProQuest, AISeL, and ScienceDirect. Content analysis of relevant articles resulted in an ontology comprising 131 classes of information deemed relevant to be addressed in privacy notices by extant research.	Establishing TIPP requires access to a diverse and comprehensive set of information on privacy practices.	(Dehling 2017, Dehling et al. 2014)

#	Idea	Study	Findings leading to rejection of idea	Key implications for TIPP theory	Reference
2	Privacy notices are useful artifacts to establish TIPP.	Content analysis of privacy notices of top 300 widely used mobile health (mHealth) smartphone applications (apps) on Android and iOS.	Privacy notice not available for 69.5% of analyzed apps Privacy notices are too long (1,755 words on average). Privacy notices are too hard to comprehend (average reading grade level of 16 years of education). Two thirds 66.1% of privacy notices did not focus on the corresponding app.	Design of transparency artifacts has to account for the cognitive capabilities of consumers. App providers often fail in offering useful content in privacy notices.	(Sunyaev et al. 2015)
3	Establishing TIPP is irrelevant because many IS do not pose privacy risks.	Cluster analysis of privacy risks introduced by 2,452 Android and 21,953 iOS mHealth apps based on content analysis of information available from official app stores.	Identification of 245 distinct clusters grouped into 12 archetypes of apps with similar potential damage for consumers due to privacy risks. Only 4.37% of apps in sample posed no discernable privacy risks. Archetype with highest potential for damage due to privacy risks comprised 11.67% of apps in sample.	Privacy risks posed by IS are diverse and depend on functionality offered by IS. Most relevant information for transparency artifacts depends on the privacy risks of the IS for which TIPP should be established.	(Dehling et al. 2015)
4	The main issue preventing establishment of TIPP are challenges in user interface design.	Development of an artifact for comparing privacy risks between mHealth smartphone apps (based on datasets from idea #2 and idea #3).	Once information on privacy practices is available, privacy risks can be compared based on a normalized risk score. More details on risks can be offered by making assessment results for each risk factor available on demand. Differences in consumer needs can be represented by allowing consumers to tailor risk score calculation by adapting factor weights. Privacy risk assessment needs to evolve in line with new privacy practices identified in apps.	Frontend development is not the main challenge for emergence of useful transparency artifacts. A more pressing challenge is access to reliable and comprehensive information on privacy practices and consumers' current information needs.	(Brüggemann et al. 2016)

#	Idea	Study	Findings leading to rejection of idea	Key implications for TIPP theory	Reference
5	Tracking of privacy practices in an IS can be automated.	Development of a static code analysis pipeline to obtain information on relevant privacy practices (drawn from literature review from idea #1) in Android apps (n=317). Comparison of automated with human review for a subset of apps (n=6).	Only information on privacy practices carried out in the IS can be obtained. Static code analysis is not useful to obtain information on privacy practices not carried out. Many Android app binaries are not easy to obtain on scale. Many app providers impede static code analysis by employing source code obfuscation, which reduces the accuracy of analyses. Static code analysis outperforms human reviewers in terms of speed, cost, and consistency, but human reviewers can better interpret source code context and find more detailed information on privacy practices.	Tracking of privacy practices can be automated to some degree. Comprehensive information collection on privacy practices requires multiple information sources. Automated review is useful for fast analysis of many apps and app versions, while human review is useful to obtain more detailed information due to more focused assessments.	(Brüggemann et al. 2019)
6	Consumers have homogeneous needs for information on privacy practices.	Online survey of consumer needs for information on 31 classes of information relevant to privacy notices (drawn from literature review from idea #1).	Cluster analysis of participant (n=134) responses revealed 10 clusters with different information needs ranging from low in all classes to high in all classes of information. Unpublished similar follow-up survey (n=909) largely replicated the findings and allowed for exploratory factor analysis grouping information needs into five factors: (1) how information is collected, (2) collection of sensitive information, (3) collection of information about consumers, (4) how information is used, and (5) available privacy controls.	Consumer needs for information on privacy practices do not exhibit strong regularities; thus, they are not easily predictable. Transparency artifacts must be adaptive to different information needs of consumers because information needs are too diverse and dynamic to be usefully served in a predetermined way.	(Dehling et al. 2016)

REFERENCES

Ahmad W, Clark DD (2021) A systems approach toward addressing anonymous abuses: Technical and policy considerations. *IEEE Security & Privacy* 19(2):38–47.

Altman I (1977) Privacy regulation: Culturally universal or culturally specific? *Journal of Social Issues* 33(3):66–84.

Brüggemann T, Dehling T, Sunyaev A (2019) No risk, more fun! Automating breach of confidentiality risk assessment for Android mobile health applications. *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS 2019)*. (Wailea, Maui, HI, USA), 4266–4275.

Brüggemann T, Hansen J, Dehling T, Sunyaev A (2016) An information privacy risk index for mHealth apps. Schiffner S, Serna J, Ikonomou D, Rannenberg K, eds. *Proceedings of the 4th Annual Privacy Forum*. (Springer International Publishing, Frankfurt (Main), Germany), 190–201.

Cavoukian A (2009) Privacy by design: The 7 foundational principles. *Information and privacy commissioner of Ontario, Canada*. Retrieved (January 13, 2023), https://privacy.ucsc.edu/resources/privacy-by-design---foundational-principles.pdf.

Day G, Stemler A (2019) Infracompetitive privacy. *Iowa Law Review* 105(1):61–106.

Dehling T (2017) RECIPE: An ontology of the information relevant for organizational information privacy communication. *University of Kassel Working Paper Series*. (Kassel, Germany).

Dehling T, Gao F, Schneider S, Sunyaev A (2015) Exploring the far side of mobile health: Information security and privacy of mobile health applications on iOS and Android. *JMIR mHealth and uHealth* 3(1):e8.

Dehling T, Gao F, Sunyaev A (2014) Assessment instrument for privacy policy content: Design and evaluation of PPC. *Proceedings of the Pre-ICIS Workshop on Information Security and Privacy*. (AIS, Auckland, New Zealand).

Dehling T, Schmidt-Kraepelin M, Demircan M, Szefer J, Sunyaev A (2016) User archetypes for effective information privacy communication. *Proceedings of the Pre-ICIS Workshop on Information Security and Privacy*. (AIS, Dublin, Ireland).

Milne GR, Pettinico G, Hajjat FM, Markos E (2017) Information sensitivity typology: Mapping the degree and type of risk consumers perceive in personal data sharing. *Journal of Consumer Affairs* 51(1):133–161.

Mulligan DK, Koopman C, Doty N (2016) Privacy is an essentially contested concept: A multi-dimensional analytic for mapping privacy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374(2083):20160118.

OECD (1980) OECD guidelines on the protection of privacy and transborder flows of personal data. *Recommendation by the Council of the OECD*. Retrieved (January 13, 2023), http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm Archived at: http://www.webcitation.org/6Y7eE1Q3m.

Peppet SR (2011) Unraveling privacy: The personal prospectus and the threat of a full-disclosure future. *Northwestern University Law Review* 105(3):1153–1204.

Rubinstein IS, Good N (2013) Privacy by design: A counterfactual analysis of Google and Facebook privacy incidents. *Berkeley Technology Law Journal* 28(2):1333–1413.

Solove DJ (2002) Conceptualizing privacy. *California Law Review* 90(4):1087–1155.

Spiekermann S, Cranor LF (2009) Engineering privacy. *IEEE Transactions on Software Engineering* 35(1):67–82.

Sunyaev A, Dehling T, Taylor PL, Mandl KD (2015) Availability and quality of mobile health app privacy policies. *Journal of the American Medical Informatics Association* 22(e1):e28–e33.

US Federal Department of Health Education and Welfare (1973) Records, computers and the rights of citizens: Report of the secretary's advisory committee on automated personal data systems. Chapter III. Safeguards for privacy. Retrieved (January 13, 2023), https://epic.org/privacy/hew1973report/c3.htm.

Westin AF (1967) *Privacy and freedom* (Ig Publishing, New York, NY, USA).