

Comments concerning  
National Privacy Research Strategy  
(in response to RFI by NSF of 09/18/14)

Dusko Pavlovic  
University of Hawaii  
dusko@hawaii.edu

## 1 Introduction and overview

Privacy is not just a security requirement, but first of all a *right* established through social contract, and balanced through policy decisions. Security research should therefore not be expected to resolve all privacy problems by technical means, since most of them are not (merely) technical problems, but involve social and political processes. A suitable role for security research should be to inform the stakeholders in the privacy policy debates, to provide solid foundations for the privacy decisions, and to develop the needed tools for implementing privacy policies and privacy decisions.

The public discourse about privacy is hampered by several conceptual difficulties. One is that the fundamental idea of privacy as a basic constituent of social life is often replaced by narrow views, influenced by the latest news about the privacy threats arising from the latest communication technologies. This difficulty and the research needed to resolve it is commented about in Sec. 2. The second difficulty is that similar privacy problems, arising in different areas, are often not recognized as similar, but conceptualized differently in each area, and the privacy solution in one area is treated as the privacy problem in another one, depending on the political and economic influence of the stakeholders. This difficulty and the research needed to mitigate it are further explained in Sec. 3. Sec. 4 describes an emerging path towards balancing privacy. Sec. 5 discusses an approach to assessing and measuring privacy, that may also be useful in specifying and designing privacy policies, protocols and mechanisms.

## 2 Why is it hard to agree about privacy objectives?

The ongoing public discussions about privacy have been driven by the privacy threats arising from the new technologies ever since Warren and Brandeis' seminal paper [13] from 1890. The definition of privacy as the *right to be left alone* goes back to that paper. The phrase 'being left alone' here does not only mean being left alone by the the government officials, but also by the journalists and by the traveling salesmen; and not only alone with your family and your tangible properties, but also with your intangible properties, your inventions, culture and beliefs. Following the subsequent expansion of the technologies that shifted the bounds of privacy in various ways, the public perception of privacy has mainly been focused on the information gathering and the information flow control techniques of the day, gradually losing sight of private life as a basic component of social life.

Without the common ground of this overarching view, the participants of the privacy debate have been advocating different notions of privacy, each serving a particular social goal that the participant who advocates it sees as central. The task of reconnecting these particular views of privacy within a common conceptual framework, fine tuning the demarkation line between the private and the public sphere requires a genuinely interdisciplinary discourse, unifying social sciences and computer science as effectively as social networks have been unified with computer networks.

In contrast with the other main topics of security research (such as the properties from the confidentiality-integrity-availability triad), privacy is not just a design requirement that we impose as desirable on a system or on a process. As explained by Warren and Brandeis [op. cit], privacy is a *right*, a political of legal entitlement of a participant in a social process. The right to privacy is viewed as an expansion of the right to life; the ownership of private data is viewed as an expansion of the ownership of physical assets. Much earlier, Aristotle [1] explained that the fundamental constituent components of a society are always

- the private sphere, comprising home, family, childbirth, etc., and
- the public sphere, comprising city, market, war, etc.

The balance between the two spheres changed from state to state, and from war to war. The conflict between the two spheres was depicted, e.g., in Sophocles' tragedy *Antigone* [12], where the heroine was torn between her private duty to bury her brother who perished in a war against the king, and her public duty to obey king's order to leave her brother unburied in the field.

Aristotle's subdivision into the private sphere of home (οἶκος, which is the root of our words *economy* and *economics*) and the public sphere of the city (πόλις, which is the root of our words *policy* and *politics*) is echoed in the modern subdivision of the methods of social choice into the economic mechanisms of the market, and the political mechanisms of voting and representation [2]. The problem of demarkation between the public sphere and the private sphere, addressed by the diverse social structures proposed and tested throughout history, is nowadays expressed through dynamics of the economic and the political forces, and their efforts to control each other. Each side thus sees the necessity of their social role as the foundation of freedom, and the defining factor of privacy, leading them to disagree about the very definition of the concept under discussion. In some cases, one side emphasizes the need for access to information, the other one the need for compartmentalization of that information; in other cases they switch the roles and each emphasizes the opposite concern with respect to the other type of data.

The answer that each society offers at the end of the day to the question of privacy is decided at the deepest layers of its political and economic system, and determined by the real political and economic powers. The researchers can provide the terms that can be used to state that answer, and they can propose the tools to implement it. They can also dispel some misunderstandings.

### 3 Privacy by any other name

The privacy requirements usually take the form of imposing precise boundaries on the information flows of specific type. E.g., my medical files should be available to my physician, but not to my employer. The conflicts arise when there are opposite reasons, one requiring that an information is available, the other one that it is restricted. E.g., a wiretap of a phone call between a known criminal and his mother should be available to an investigator because the criminal may be involved

in a criminal activity, but it should be restricted to preserve the rights of his mother, if she is not a suspect. The other way around, the information about the wiretap should be available for scrutiny, to prevent the abuses of policing powers, and it should also be restricted, to prevent that the suspects avoid surveillance. To resolve both pairs of logically contradictory requirements, a privacy policies must go beyond logical specifications, and specify a social preference providing a balanced decision for each case, and finely tuned information flow controls to implement both decisions.

Technically, the implementations of individual privacy policies thus boil down to finely tuned information flow controls. But the implementations of intellectual property and digital rights policies also boil down to finely tuned information flow controls. And the implementations of the policies for disclosure of classified data also boil down to finely tuned information flow controls. The same technical task of restricting the public flows of information is studied under three different names for three different types of data ownership: as privacy for the individual data, as digital rights for the commercial data, and as classification for the government data. The same agents often work on protecting the information flow restrictions in one domain, and on circumventing or breaking such restrictions in another domain. The similarity of the tasks and the differences of the results and of the incentives between the domains of digital rights and individual privacy were commented about in [15, 6]. To further complicate things, the three research directions, sponsored by the three families of stakeholders, with their different funding powers, lead to three different families of techniques for information flow controls on one hand, and for information gathering and circumventing the flow controls on the other hand. Fig. 1 displays a crude picture of the effects.

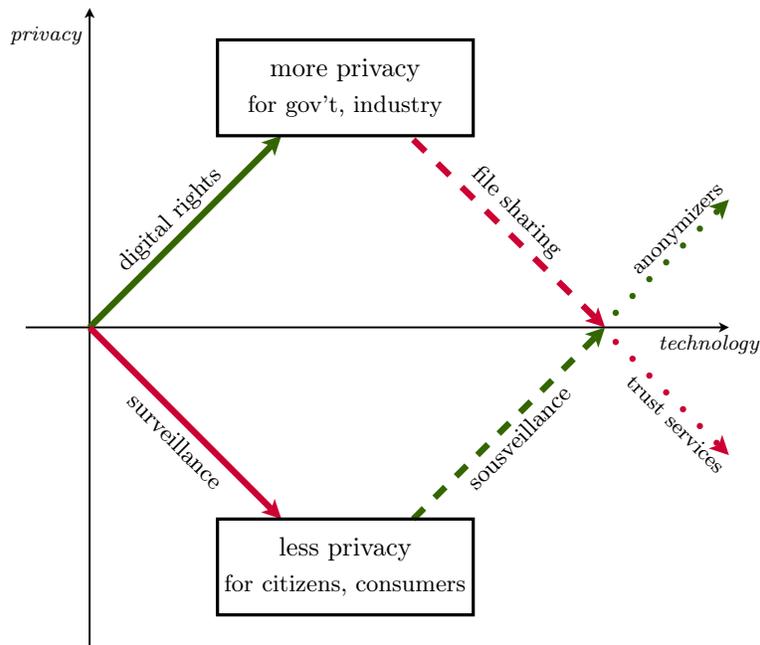


Figure 1: Balancing privacy through information gathering and information flow control

Although the differences between the applications in some cases hamper the direct transfers of the techniques from one domain to another (e.g., the DVD encoding techniques do not yield the protections for private databases in any obvious way), in other cases the application gap between the

domains seems smaller. E.g. the definitions and some of the methods of differential privacy [4] seem applicable not only to statistical databases, but also to the disclosure mechanisms for confidential data; and the ubiquity of small cameras has led both to the increase of public surveillance and decrease of individual privacy on one hand, but also to the increase of the counter-surveillance of the public officials by private individuals, and thus to providing the tools for limiting the intrusions into the private domain. — A unified research into the underlying methodologies for information gathering and information flow controls, as applicable to all types of data and transactions, might quickly recover some of the missed opportunities for transferring the results and the techniques between the problem areas of privacy, digital rights, and classified data protections.

## 4 Privacy architectures: from mutual surveillance to mutual trust

Most paths to balancing privacy policies seem to lead through a phase of public scrutiny of the public functions that interface private data. The processes that led to balancing the controls of the private information flows through the expanse of the health care channels, or through the experiments that involve human subjects, or even through the web tracking datasets, all went through that phase. A subtle example that seems to be rapidly evolving at the moment are the processes of *'guarding the guardians'*, or *sousveillance* [7], whereby the surveillance by one participant of a social interaction (e.g. police recording a demonstration, or a company using CCTV) is balanced by the surveillance by the other participant (the demonstrators, resp. the employees or customers). The wide availability of the personal recording devices (smartphones, Google glass), combined with the expanding capability to crowdsource the evidence gathering, processing and dissemination (through various social networks) has led to spontaneous expansion of this approach as a vehicle of diverse social actions, and as a new interface between the public and the private sphere. This phenomenon requires genuinely interdisciplinary research, since the forces that drive it essentially depend on both the social and the computational preconditions, which cannot be separated.

The salient point of modeling the guarding-the-guardians processes, as a process of balancing the public functions and the privacy requirements, is that already the simple network models (along the lines of [8, 9]) uncover an interesting and promising social dynamics — whereby the expansion of the private surveillance of public functions in response to the increase of public surveillance of private actions does not lead to an escalation of conflict, or even to a surveillance arms race, as one might expect at the first sight, but to an actual *decrease* of surveillance on both sides, by introducing *mutual trust*, arising from the availability to the protocol participants of information flows about the ongoing protocol executions [9]. By providing grounds for trust, the open mutual surveillance can thus decrease the need for surveillance. This negative feedback seems worth attention and research.<sup>1</sup>

## 5 Privacy assessment and measurement

If the privacy breaches are viewed as information leaks, then analyzing privacy requires recognizing and quantifying the flows of private information. But while information theory provides the mathematical tools for quantifying the amount of information flowing through a channel, taking computational feasibility into account leads to a slew of negative results, some immediate, some

---

<sup>1</sup>The other way around, the same toy toy model suggests that one-sided surveillance generate an exponential (thus self-destabilizing) expansion of surveillance. But this phenomenon is already known from the historic evidence about the totalitarian regimes.

nontrivial, showing that the various simple and natural requirements that one might want to impose on the databases containing private data (cf. [3]) cannot be feasibly realized [4]. The theory of *differential privacy* [5] addresses the algorithmic aspect of the problem.

The focus on algorithmics, however, seems to have impacted the practical applicability of the approach, or at least slowed down the actual applications. The definition of differential privacy provides a framework for specifying algorithms that assure that the information leakage negligible in a formal sense agreed by the researchers. In the practical applications, though, the notion of acceptable or negligible leakage varies from case to case. While a robust notion of negligible information leakage is convenient for formulating theoretical results, the practical tasks of assessing the privacy loss from application to application would be better served by a flexible and more informative measurement tool for quantifying the information leakage, which would allow the practitioners to gather experience from case to case. Such measurement tools are not far from the conceptual realm of differential privacy, and the research towards spelling them out might quickly lead to a tide of applications. Looking at the problem through the lens of our own work within the ongoing projects, we believe that the tools for quantifying privacy can be obtained by combining the theory of differential privacy with the structure of proximity sets, previously used for concept mining and trust network analysis [10, 11].

## 6 Conclusion

The privacy objectives, and the demarcation between the public sphere and the private sphere is determined by balancing the social powers of the involved social agents and interest groups. What will be public and what private is determined through social contract. Totalitarian regimes deny all privacy and intellectual property, while covering their own actions by shroud of secrecy. In commercially driven societies, the private and the government data are freely collected and sold for commercial purposes, while the strong protections are only provided for the commercial data and intellectual property. In some early societies, like pre-Doomsday Book England, the public sphere boils down to the royal court and the market place, and even the very existence of the individuals living in the country was their private matter, until the need for tax collections led to the first data collections. Even nowadays, some extreme advocates of privacy endorse the old Saxon idea of the local sheriff as the highest public office, beyond which there is no higher political power [14]. The extreme views of the private and of the public domain span from one end of the political spectrum to the other. They illustrate the scope of the misunderstandings to which the conceptual compartmentalization of the public debate about privacy can lead — and the urgency of a conceptual unification, providing a common ground for a productive social discourse and technical innovation in the realm of privacy.

## References

- [1] Aristotle. *The Politics*. Penguin, 1981.
- [2] Kenneth J. Arrow. *Social Choice and Individual Values*. Cowles Foundation monograph. Yale University Press, 1963.
- [3] Tore Dalenius. Towards a methodology for statistical disclosure control. *Statistik Tidskrift*, 15(429-444):2–1, 1977.

- [4] Cynthia Dwork. Differential privacy. In Michele Bugliesi et al., editor, *Proceedings of ICALP 2006, Part II*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2006.
- [5] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3-4):211–407, 2014.
- [6] Lawrence Lessig. The code of privacy. *Proc. of Am. Philosophical Soc.*, 151(3):283–290, 2007.
- [7] Steven Mann. Veilance and reciprocal transparency: Surveillance versus sousveillance. In *Proceedings of ISTAS 2013*, pages 1–12. IEEE, 2013.
- [8] Dusko Pavlovic. Dynamics, robustness and fragility of trust. In Pierpaolo Degano, Joshua Guttman, and Fabio Martinelli, editors, *Proceedings of FAST 2008*, volume 5491 of *Lecture Notes in Computer Science*, pages 97–113. Springer Verlag, 2008. arxiv.org:0808.0732.
- [9] Dusko Pavlovic. Quantifying and qualifying trust: Spectral decomposition of trust networks. In Pierpaolo Degano, Sandro Etalle, and Joshua Guttman, editors, *Proceedings of FAST 2010*, volume 6561 of *Lecture Notes in Computer Science*, pages 1–17. Springer Verlag, 2011. arxiv.org:1011.5696.
- [10] Dusko Pavlovic. Quantitative Concept Analysis. In Florent Domenach, Dmitry I. Ignatov, and Jonas Poelmans, editors, *Proceedings of ICFCA 2012*, volume 7278 of *Lecture Notes in Artificial Intelligence*, pages 260–277. Springer Verlag, 2012. arXiv:1204.5802.
- [11] Dusko Pavlovic. Bicompletions of distance matrices. In Bob Coecke, Luke Ong, and Prakash Panangaden, editors, *Computation, Logic, Games and Quantum Foundations. The Many Facets of Samson Abramsky*, volume 7860 of *Lecture Notes in Computer Science*, pages 291–310. Springer Verlag, 2013.
- [12] Sophocles. Antigone. [classics.mit.edu/Sophocles/antigone.html](http://classics.mit.edu/Sophocles/antigone.html).
- [13] Samuel D. Warren and Louis D. Brandeis. The right to privacy. *Harvard Law Review*, 4(5):193, 1890.
- [14] Wikipedia. Sovereign citizen movement. [en.wikipedia.org/wiki/Sovereign\\_citizen\\_movement](http://en.wikipedia.org/wiki/Sovereign_citizen_movement).
- [15] Jonathan L. Zittrain. What the publisher can teach the patient: Intellectual property and privacy in an era of trusted privity. *Stanford Law Review*, 1201, 2000.