Is Privacy Regulation Slowing Down Research on Pervasive Computing?

Claudio Bettini – *Università degli Studi di Milano*, Milan, Italy
Salil Kanhere – *University of New South Wales*, Sydney, Australia
Marc Langheinrich – *Università della Svizzera italiana (USI)*, Lugano, Switzerland
Archan Misra – *Singapore Management University*, Singapore
Delphine Reinhardt – *University of Göttingen*, Göttingen, Germany

***Abstract*—Privacy legislation has often been identified as a roadblock for advanced context-aware pervasive applications. We gathered feedback from over 150 researchers in pervasive computing in order to better understand if this attitude is still shared. Our findings indicate that most respondents do not feel any major impediments in adhering to privacy regulations, and also are apparently familiar with the latest legal developments. Has pervasive computing's privacy challenge been "solved"?**

Introduction

Reading about large-scale data privacy violations has become commonplace. 2018 alone saw almost 1 billion user accounts involuntarily disclosed or hacked, including customers from restaurant chains, retailers, and hotels, as well as major online services (Facebook, Uber); in July 2019, over 2 billion log entries from IoT management platform Orvibo were stolen, containing user accounts, passwords, and even recorded "smart camera" conversations. Mobile and pervasive technology and services have a role in this scenario, since they introduce new devices and communication protocols (all with plenty of room for new vulnerabilities), new types of personal data, and a new scale for the amount of data being collected [11].

Driven in part by these large-scale privacy violations and increasing public concern, regulators in most countries have taken action in revising legal obligations related to personal data protection. In 2016, the EU approved its new *General Data Protection Regulation* (GDPR), which went into effect in May 2018; in the US, the state of California passed a new data privacy law in June 2018 that in many ways resembles the GDPR (the so-called "California Consumer Privacy Act of 2018," CCPA); Japan has taken similar steps with the amended Act on the Protection of Personal Information.

As a consequence, industry has begun to invest a significant amount of money and resources in ensuring the continuing compliance of their systems and products with the changing legal landscape, including adjusting the design, production, and test processes of their products. This will inevitably have an impact on the type of services that will be offered, on their cost, and on the timing of their appearance on the market. While the value of investing in security is usually well understood, investing in privacy is often seen only as a cost, especially by small and medium sized enterprises.

This seemingly unavoidable trade-off (more privacy means less services) in principle applies also to research in this space: while there are usually ample exceptions for research, legislation does have a significant impact once research prototypes are meant to move into commercialization. Similarly, increased privacy awareness has also heightened the bar for getting ethical clearance both at an institutional level and within the wider research community. We are not the first to investigate how privacy affects pervasive research. One of the first efforts in this space came from Langheinrich in 2003 [9], who found a high level of non-concern among researchers. While both the legal landscape and research practices have significantly changed since then, Bednar et al. [1] found similar levels of disinterest more than 15 years later while interviewing six senior software engineers. In 2019, Spiekermann et al. [16] surveyed 124 engineers to find that while most considered privacy important, few enjoyed including it within their systems. These last two studies also found that many engineers struggle with the organizational environment: They face a lack of time and autonomy that is necessary for building ethical systems. A similar effort by Szekely [17] focused on IT professionals, though their survey was targeted on surveillance issues and limited to two countries: Hungary and the Netherlands, highlighting a higher level of awareness in the second country.

Our work explicitly focuses on researchers. Specifically, we wanted to understand how well the research community in pervasive computing understands current privacy legislation, how it affects their work, and how their attitudes towards ethical decisions in this space vary. Within the context of a privacy-focused panel at the IEEE 17th International Conference on Pervasive Computing and Communications (PerCom 2019) we conducted a brief survey among active researchers in this field to inform the panelist's discussions. This article summarizes their responses and comments on their implications for the field.

I. The survey

A. Settings and Sample

Our online questionnaire including ten questions (see sidebar) was distributed through the PerCom conference mailing list. The survey took about 5 minutes to complete and no rewards were provided to the participants. 154 researchers on mobile and pervasive computing from both academic and industrial institutions responded. When asked about their personal attitude regarding the (digital) sharing of personal data, 49% indicated belonging to the category "quite concerned", who "want to know exactly who gets [their] data and what they do with it". The second most represented category was "quite liberal" with 28%, followed by "very concerned" with 18%. The remaining 5% of participants indicated to be "very liberal", i.e., "enjoying sharing to a large audience including location, pictures, video". Mapping the central two categories (quite concerned, quite liberal) to a "pragmatist" approach to privacy, these results are roughly in line with the distribution of Westin's privacy categories (pragmatist, unconcerned, fundamentalist) found in prior surveys [8]. It also means that the large majority of researchers is concerned about the protection of their personal data when acting as *users* of digital services.

In order to better situate these concerns within a concrete pervasive computing setting, we asked participants to select up to two types of sensor data (from an overlapping list of 4 examples) that they thought should *not* be collected or retained by municipal authorities in a "smart city" application: (i) their indoor location in public spaces (e.g., shopping centers); (ii) their outdoor (cellular) location data; (iii) their location data via video analytics; and (iv) scraping their publicly available social media data. Alternatively, participants could select "all of the above".

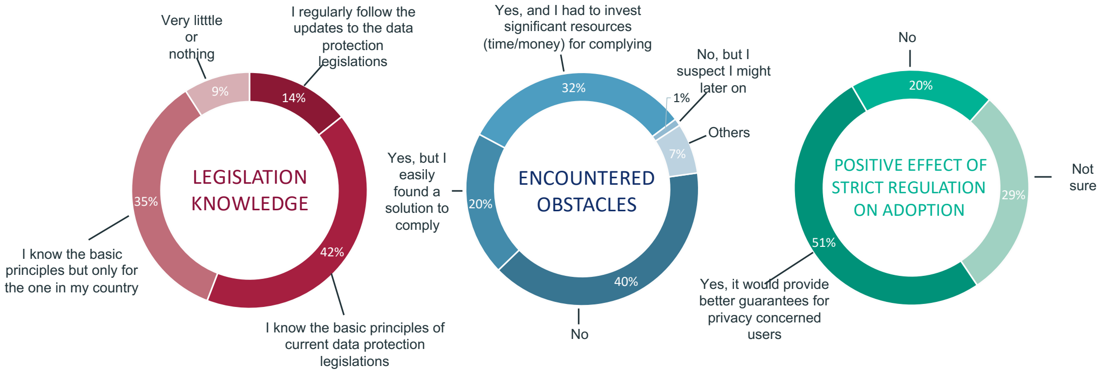

Fig. 1. Participants' knowledge, experience, and belief about privacy regulations

While about one in four participants were happy that all of this data would be collected (25%), over two thirds of our participants (69%) would not support video-based tracking. Interestingly enough, only 31% considered indoor location sharing to be problematic, even though the survey did not specify the tracking method (e.g., via BLE tags, or in fact video analytics). Outdoor (cellular) location tracking saw similar levels of concern (33%), while sharing social media data had only 18% of participants concerned. This example illustrates that the type of sensor plays an important role in people's privacy concerns: even though the end-result of, say, location-tracking via a cellular signal may be identical to a location trace obtained using video tracking (or, depending on the density of cameras, could even be much more detailed), our participants reacted strongly to the idea of video-based tracking. Also, the fact that most participants did not worry about having their social media streams monitored seems to suggest that they are either not aware of the significant risks that such a practice entails (see, e.g., [14]) or consider any and all public data "lost" from their control [6].

B. *Regulation Awareness and Knowledge*

We first asked our participants how much they estimate to know about general data protection legislation. 14% of our respondents reported that they regularly follow updates to data protection legislation. 42% indicated that they know the basic principles of current (!) data protection legislation; further 35% stated that they knew those basic principles only for the laws applying to their country. Combined, this still means that over 91% of our respondents felt that they have a good grasp on current national privacy legislation – an astonishingly high value. Only 9% indicated knowing very little or nothing about privacy law. Obviously, such self-assessment is no proof of actual knowledge, and we did not rigorously test the actual legal understanding of our participants.

However, a follow-up question then asked participants to indicate in which countries the EU GDPR would apply. Almost half of our respondents (47%) correctly answered that it applies wherever European Union citizens are served ("marketplace rule"; Article 3(2) GDPR). 36% incorrectly assumed that it would only apply within the European Union, while 9% believed the GDPR to apply for all countries in Europe. The remaining 8% believed that it would apply worldwide to all citizens. While clearly a low bar, understanding the scope requirement is an important aspect of today's legal privacy landscape, as laws increasingly use this approach. For example, California's CCPA similarly applies to data controllers that "do business in the State of California", irrespective of their physical establishment. The answers to this question seem to indicate that, in practice, there is still a gap between legislative intent and practitioner awareness.

C. System Design Approach and Privacy-by-Design

Since our sample was composed of researchers working in the area of pervasive computing, we were interested in knowing how privacy legislation has affected them in their development of pervasive computing technologies and applications. A very positive result is that a large majority of our respondents either indicated that they "always" think about privacy issues when developing new solutions (35%) or that they do so if the system "would handle very sensitive information" (52%). However, 12% would not consider privacy in the design phase, but only in a later phase when the prototype would need to be tested/deployed. 1% indicated that they never think about privacy issues.

Given the above numbers, it is not surprising that a majority of our sample (59%) agreed that the obligation imposed by the GPDR of incorporating "privacy by design" in products or systems is a good thing, as it "saves costs later for compliance and leads to better products". However, 27% of our participants indicated that they did not know what "privacy by design" meant or what it involved. The remaining 14% were against this obligation, indicating that it required "special expertise" and "would not always be needed".

D. Applying Privacy Protection and Potential Obstacles

Moving from theory to the practice of privacy protection, we were interested in knowing which privacy protection techniques our participants already implemented in their systems.

When presented with several key categories of privacy protection methods, a large group of participants (39%) reported to use only basic security measures like access control. The remaining used the following techniques: anonymization (25%), cryptography and/or blockchain-based techniques (9%), or obfuscation or statistical perturbation including differential privacy (6%). About one out of five participants (21%) declared to use more than one of the above techniques.

Among the participants having already included privacy-preserving techniques in their solutions ($n = 97$, multiple choices possible), only 21% indicated that they did not encounter any difficulties. In contrast, 50% reported that it was difficult to ensure that the implemented solution was the most efficient one. 33% found it difficult to translate and apply the legal requirements into their solution, while 30% had difficulties to choose among the different solutions to best protect privacy.

E. Impact of Regulation on Adoption

Regarding the more general issue of the impact of regulation on the design and deployment of pervasive solutions, 43% of our participants declared that they did not (yet) face obstacles due to privacy regulation. Some participants had to deal with privacy issues, but easily found a solution to make their system legally compliant (20%). However, a high number of them (32%) had to invest significant resources in terms of time and money for complying with privacy legislation. A few participants (5%) indicated that they had to previously abandon a project due to privacy legislation.

The last question of the survey collected opinions in favor of or against strong privacy laws. A majority of our participants (51%) indicated that a strict privacy regulation should have a positive effect on the adoption

Fig. 2. SIDEBAR: Survey Questions and Possible Answers

1) What is your attitude about sharing your personal digital data?
    - Very liberal (I enjoy sharing to a large audience including location, pictures, video)
    - Quite liberal
    - Quite concerned (I want to know exactly who gets my data and what they do with it)
    - Very concerned (I protect my email, mobile number, I do not post on socials, I usually deny consent)
2) Do you think about privacy issues when you design a new mobile/pervasive solution?
    - Always
    - Yes, but only if it handles very sensitive information
    - No, only later after the prototype is ready and needs to be tested/deployed
    - Never
3) How much do you know about the data protection legislation?
    - I regularly follow the updates to the data protection legislations
    - I know the basic principles of current data protection legislations
    - I know the basic principles but only for the one in my country
    - Very little or nothing
4) Do you know where (countries) the EU GDPR applies?
    - Worldwide to all citizens
    - All of Europe
    - Only EU
    - Wherever EU citizens are served
5) Which of these types of context/personal sensor data do you think applications run by CITY/MUNICIPAL authorities should NOT be authorized to collect or retain (due to privacy concerns)? (Pick at most 2)
    - Indoor location in public spaces
    - Outdoor cellular location data
    - Person identity (from video analytics) in public spaces
    - Publicly crawl-able social media data
    - They can collect all the mentioned data if they declare the purpose and I believe it is useful
6) What do you think about the obligation of incorporating ”privacy by design” in products/systems?
    - I agree, it saves later costs for compliance and leads to better products
    - I disagree, this requires special expertise and it is not always needed
    - I do not know what ”privacy by design” is or what it involves
7) Have you ever tried to include some privacy protection technique in your solutions?
    - Yes, anonymization
    - Yes, obfuscation or statistical perturbation (including differential privacy)
    - Yes, cryptography and/or blockchain based techniques
    - Yes, more than one of the techniques mentioned in other answers
    - No (possibly only basic security measures like access control)
8) If you answered Yes to the previous question, have you ever encountered difficulties when including privacy protection techniques in your solutions?
    - None, it was always straightforward
    - It was difficult to translate and apply the legal requirements into my solution
    - It was difficult to choose among the different solutions to best protect privacy
    - It was difficult to ensure that the implemented solution was the most efficient one
9) Did you experience obstacles to your work in pervasive computing due to the privacy regulation?
    - Yes, and I had to abandon my project
    - Yes, and I had to invest significant resources (time/money) for complying
    - Yes, but I easily found a solution to comply
    - No
    - No, but I suspect I might later on
    - Other (free text)
10) Do you think that a strict privacy regulation will have a positive effect on the adoption of pervasive systems?
    - Yes, it would provide better guarantees for privacy concerned users
    - No
    - Not sure

of pervasive systems, providing better guarantees for privacy-concerned users as well as for developers. A good number of participants disagreed (20%), while a third was not sure either way (29%).

II. Discussion

We see three key issues emerging from our survey: (1) the benefits of strong privacy legislation for research; (2) the lack of guidance for implementing privacy-by-design; and (3) fundamental challenges in today's privacy regimes. We will briefly discuss these in turn below.

A. Benefits of Legislation

As our participants indicated, strict regulation can have tangible benefits not only for users (who may more readily adopt a pervasive service) but also for service providers. In fact, having clear legal guidelines has been beneficial for several of the authors of this article. While it has become commonplace to obtain ethical clearance from an institutional review board prior to running a particular study, universities increasingly include their legal departments when it comes to authorizing field deployments. For example, when one of the authors (Archan) attempted to deploy smart services (see, e.g., [7]) across his university's campus, legal services were actually grateful for the concrete legal guidance offered by Singapore's 2012 privacy law (the 2012 Personal Data Protection Act, PDPA). Having concrete rules and practical guidelines will be essential in order to ensure that privacy laws reduce uncertainty for researchers, rather than increasing it (see also II-B). Similarly, if achieving legal compliance can only be possible if service quality is reduced (e.g., by removing or limiting a system's personalization capabilities), users may not perceive any benefits of such legal protection and simply "vote with their feet" by using more complete but less privacy-friendly systems. It might be time to reconsider the principal approach to privacy legislation, which in many jurisdictions by default attempts to minimize the collection and procession of personal data (see also II-C).

B. Privacy-By-Design Guidance

While less than a third of our participants indicated that they did not know what "privacy by design" meant, this still is a significant number, especially since our respondents can all be considered technology experts. This points to the obvious gap between the GDPR's well-meant inclusion of this principle (Article 25: "Data protection by design and by default") and the lack of concrete technical guidance on how to implement this principle. Specific privacy enhancing technologies for mobile and pervasive computing, often inspired by solutions in databases, have been investigated for almost two decades leading to a very rich set of methods – recent surveys can be found in Bettini et al. [2], Christin [3], and Gkoulalas-Divanis et al. [5]. However, this wealth of methods does not mean that it is any easier in applying the right method in the right context. Researchers need more guidance on how to incorporate these technologies and procedures in their system, already at design time. While several attempts on formulating more concrete methodologies in this space have been made (see, e.g., Chapter 5 in [12]), the huge variety of pervasive computing systems render the idea of a simple "how-to" that could be followed in each and every project infeasible. Instead of refining and extending our vast array of methods for protecting personal data, we need more research into the practical application (i.e., integration) of such techniques into pervasive systems.

C. Fundamental Challenges

About one third of our respondents indicated they had to "invest significant resources for complying" or even had to abandon their project due to the challenges posed by privacy laws. Legal scholars have long since challenged the suitability of current privacy laws for today's technology landscape. Even though both the GDPR and the CCPA were created with social media firmly in mind, they nevertheless still trace their roots back to the privacy laws of the 1970s – when data was still stored on punch cards. Not only was the amount of stored and processed data minuscule compared to today's volume (according to the World Economic Forum, the amount of digitally stored data will reach 44 Zetabytes in 2020 – that's 40 times more bytes than stars in the universe [4]), the key stakeholders were predominantly governments, and most of the captured data was still manually entered. Among the key criticisms are the predominant focus on personal rights versus overarching social benefits [13]; the belief that data can actually be anonymized (and that it matters) [10]; and the fundamental limits of meaningful notice and choice [15]. Veil [18] in particular criticizes the GDPR's "one-size-fits-all" approach, which may make it easy for lawmakers to craft legislation but which imposes much of the same obligations (Veil counts no less than 68!) on both large international companies (irrespective of them operating social media sites or, say, manufacturing escalators), a plumber, or a small church club. While large legal frameworks that apply across many legal contexts are appealing, the vast differences in data collection motivations and corresponding privacy risks may require a much more individualized approach.

III. Conclusions

The main goal of our study was to investigate, discuss, and better understand the impact of the recent evolution of personal data protection legislation on the work of the research community on mobile and pervasive computing, and possibly identify critical issues that deserve more attention by researchers and/or by regulators.

Among the several insights, three partially unexpected key observations emerge from our survey:

- Most of the researchers believe that they are well-aware of their legal obligations with respect to privacy;
- Only very few researchers mention insurmountable obstacles due to privacy regulation – the majority of respondents either did not face obstacles or found a privacy-preserving solution;
- A majority of researchers seems to be in favor of clear and strict privacy regulation.

While we believe that our survey population is a sufficiently large and diverse sample of the target research community, we are also well aware of the limitations of our study: a) the sample is not uniformly distributed over geographical areas (slightly skewed to EU and Asia), and b) it includes mostly academic researchers. Regarding this last aspect, our results can be complemented by the results of related studies focused on the engineers and IT professionals population [1], [16], [17] as briefly discussed in the introduction.

Overall, the observations arising from our study are surprising and certainly deserve further investigation to understand, for example, the degree to which researchers are actually informed about privacy regulation applying to their specific research, and how compliant their assumed solutions are. Our study also highlights the need for enhancing the interface between privacy regulation and privacy-preserving solutions in terms of both publicly available case studies and application-specific practical guidelines. This goal can be achieved only by a multidisciplinary joint effort including legal, technical and social expertise.

References

[1] Bednar, K., Spiekermann-Hoff, S., and Langheinrich, M. Engineering Privacy by Design: Are engineers ready to live up to the challenge? *The Information Society 35*, 3 (2018), 34.

[2] BETTINI, C., AND RIBONI, D. Privacy protection in pervasive systems: State of the art and technical challenges. *Pervasive and Mobile Computing 17* (2015), 159–174.

[3] CHRISTIN, D. Privacy in Mobile Participatory Sensing: Current Trends and Future Challenges. *Journal of Systems and Software (JSS) 116* (2016), 57–68.

[4] DESJARDINS, J. How much data is generated each day? *World Economic Forum Website* (Apr. 2019).

[5] GKOULALAS-DIVANIS, A., AND BETTINI, C., Eds. *Handbook of Mobile Data Privacy*. Springer, 2018.

[6] HARGITTAI, E., AND MARWICK, A. "What Can I Really Do?" Explaining the Privacy Paradox with Online Apathy. *International Journal of Communication (19328036) 10* (2016), 3737–3757.

[7] KANDAPPU, T., MISRA, A., CHENG, S.-F., JAIMAN, N., TANDRIANSYAH, R., CHEN, C., LAU, H. C., CHANDER, D., AND DASGUPTA, K. Campus-scale mobile crowd-tasking: Deployment &#38; behavioral insights. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (New York, NY, USA, 2016), CSCW '16, ACM, pp. 800–812.

[8] KUMARAGURU, P., AND CRANOR, L. F. Privacy Indexes: A Survey of Westin's Studies. Technical report CMU-ISRI-05-138, Carnegie Mellon University, Pittsburgh, PA, USA, 2005.

[9] LAHLOU, S., LANGHEINRICH, M., AND RÖCKER, C. Privacy and trust issues with invisible computers. *Communications of the ACM 48*, 3 (Mar. 2005), 59–60.

[10] LANE, J., STODDEN, V., BENDER, S., AND NISSENBAUM, H. F. *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. Cambridge University Press, New York, NY, USA, 2015.

[11] LANGHEINRICH, M. Privacy by Design — Principles of Privacy-Aware Ubiquitous Systems. In *Proceedings of the Third International Conference on Ubiquitous Computing (UbiComp 2001)*, G. D. Abowd, B. Brumitt, and S. A. Shafer, Eds., vol. 2201 of *LNCS*. Springer, Berlin Heidelberg New York, Sept. 2001, pp. 273–291.

[12] LANGHEINRICH, M., AND SCHAUB, F. *Privacy in Mobile and Pervasive Computing*. Synthesis Lectures on Mobile and Pervasive Computing. Morgan & Claypool Publishers, 2018.

[13] NISSENBAUM, H. F. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, Stanford, CA, 2009.

[14] ROSENBLUM, D. What anyone can know: The privacy risks of social networking sites. *IEEE Security & Privacy 5*, 03 (May 2007), 40–49.

[15] SOLOVE, D. J. Privacy self-management and the consent dilemma. *Harvard Law Review 126*, 7 (2013), 1880–1903.

[16] SPIEKERMANN, S., KORUNOVSKA, J., AND LANGHEINRICH, M. Inside the Organization: Why Privacy and Security Engineering Is a Challenge for Engineers. *Proceedings of the IEEE 107*, 3 (Mar. 2019), 600–615.

[17] SZEKELY, I. What Do IT Professionals Think About Surveillance? In *Internet and Surveillance. The Challenge of Web 2.0 and Social Media.*, C. Fuchs, K. Boersma, A. Albrechtslund, and M. Sandoval, Eds. Routledge, 2011, ch. 10, pp. 198–219.

[18] VEIL, W. The GDPR: The Emperor's New Clothes - On the Structural Shortcomings of Both the Old and the New Data Protection Law. *Neue Zeitschrift für Verwaltungsrecht 2018*, 10 (2018), 686–696.

ABOUT THE AUTHORS

Claudio Bettini is full professor in the Computer Science department at Università degli Studi di Milano, where he leads the EveryWare laboratory. He received his PhD in Computer Science from the University of Milan in 1993. He has been for more than a decade, an affiliate research professor at the Center for Secure Information Systems at George Mason University, VA. His research interests cover the areas of data management in mobile and pervasive computing, data privacy and security, context-awareness and context reasoning, temporal and spatio-temporal data management. He acted as PI and co-PI of research projects on data privacy both in the US and in the EU. He is a member of the steering committee of the IEEE PerCom conference and he has been associate editor of the Pervasive and Mobile Computing Journal, The VLDB Journal, and the IEEE Transactions on Knowledge and Data Engineering. He is an IEEE senior member.

Salil Kanhere received his M.S. and Ph.D. degrees, both in Electrical Engineering from Drexel University, Philadelphia, USA. He is currently a Professor in the School of Computer Science and Engineering at UNSW Sydney, Australia. His research interests include Internet of Things, pervasive computing, blockchain, crowdsourcing, data analytics, privacy and security. Salil regularly serves on the organizing committee of a number of IEEE and ACM international conferences. He is on the Editorial Board of Elsevier's Pervasive and Mobile Computing and Computer Communications. Salil is a Senior Member of both the IEEE and the ACM.

Marc Langheinrich is full professor in the Faculty of Informatics at the Università della Svizzera Italiana (USI) in Lugano, Switzerland. His research focuses on privacy in mobile and pervasive computing systems, in particular with a view towards social compatibility. Marc is a member of the Steering Committee of the UbiComp conference series, and chairs the IoT conference Steering Committee. He has been a General Chair or Program Chair of most major conferences in the field, including Ubicomp, PerCom, Pervasive, and the IoT conference, and currently serves as the Editor-in-Chief for IEEE Pervasive Magazine. Marc holds a Ph.D. from ETH Zürich, Switzerland. He can be reached at langheinrich@ieee.org.

Archan Misra is Professor, and the Associate Dean of Research, in the School of Information Systems at Singapore Management University (SMU). Archan holds a Ph.D. in Electrical and Computer Engineering from the University of Maryland at College Park in May 2000. Over a 20-year research career spanning both academics and industry (at IBM Research and Bellcore), Archan has worked extensively on problems spanning wireless networking, mobile & pervasive computing and urban sensing. He currently directs SMU's Center for Applied Smart-Nation Analytics (CASA), which cooperates with public agencies in implementing advanced Smart Nation services in Singapore. Archan chaired the IEEE Computer Society's Technical Committee on Computer Communications (TCCC) from 2005-2007.

Delphine Reinhardt is a full professor and head of the Computer Security and Privacy group at the University of Göttingen. Delphine completed her doctoral degree in computer science (with distinction) on privacy in participatory sensing in 2013 at Technische Universität Darmstadt. Her dissertation was awarded by the Communication and Distributed Systems Group (KuVS) supported by the German Informatics Society (GI) and ITG-VDE, the Information Technology Society (ITG) of the German Association for Electrical, Electronic and Information Technologies (VDE), as well as the Association of the Friends of the Technische Universität Darmstadt for outstanding academic achievements. Her current research interests include privacy, usability, and ubiquitous computing.