International Workshop on Privacy Engineering

http://ieee-security.org/TC/SPW2016/IWPE

May 26th, 2016 at the Fairmont, San Jose, CA

Co-located with 37th IEEE Symposium on Security and Privacy

08:45-09:00	**Welcome, introductions and opening remarks**
09:00-09:15	**Privacy Engineering: Shaping an Emerging Field of Research and Practice** Addressing privacy and data protection systematically throughout the process of engineering information systems is a daunting task. Although the research community has made significant progress in theory and in labs, meltdowns in recent years suggest that we're still struggling to address systemic privacy issues. Privacy engineering, an emerging field, responds to this gap between research and practice. It focuses on designing, implementing, adapting, and evaluating theories, methods, techniques, and tools to systematically capture and address privacy issues in the development of sociotechnical systems. In this short session, Seda Gürses and Jose M. del Alamo will introduce an early definition of privacy engineering, which was elaborated taking into account the insights gathered during IWPE'15 and has been published in the IEEE Security and Privacy Magazine. ***Seda Gürses*** *is a postdoctoral research associate at Princeton University's Center for Information Technology Policy and an FWO (Fonds etenschappelijk Onderzoek–Vlaanderen) fellow at COSIC, University of Leuven. She works on privacy and requirements engineering, privacy enhancing technologies, and surveillance. Seda chairs the IWPE'16 Program Committee.* ***Jose M. del Alamo*** *is an Associate Professor in the Information and Communications Technology (ICT) Systems Engineering Department at the Universidad Politécnica de Madrid. His research focuses on personal data management issues, including privacy and identity management, in the context of software and systems engineering. Jose is the IWPE'16 General Chair.*
09:15–10:15	**Privacy and Algorithmic Accountability: Theory and Practice** Invited talk by Anupam Datta (Carnegie Mellon University) Big data analytics presents threats to privacy and related values like fairness. Our position is that accountability is a key component of the solution space. Over the last decade, my research group has developed foundations and tools for protecting privacy via accountability and extensive case studies to validate them. In this talk, I will focus on the theory and practice of algorithmic accountability. Our work is driven by the following question: When algorithmic systems, based on machine learning and related statistical methods, drive decision-making, how can we detect violations, explain decisions, hold entities in the decision-making chain accountable, and institute corrective measures? I will present two recent results in this space. First, I will describe our work on detection of violations. We have developed the first statistically rigorous methodology for information flow experiments (IFE) to discover personal data use by black-box Web services. Our AdFisher tool implements an augmented version of this methodology to enable discovery of causal effects at scale. Its application resulted in the first study to demonstrate statistically significant evidence of discrimination in online behavioral advertising, more specifically, gender-based discrimination in the targeting of job-related ads. This methodology and class of tools can be used to provide external oversight of big data systems by researchers, regulatory agencies, investigative journalists, and civil liberties groups.

	Second, I will describe our work on algorithmic transparency aimed at explaining decisions by big data systems with machine learning components. We develop a suite of quantitative input influence (QII) measures that quantify the causal influence of features (e.g., gender, age) on decisions made by a big data system. The QII measures form the basis of transparency reports that explain decisions about individuals (e.g., identifying features that were influential in a specific credit or insurance decision) and groups (e.g., identifying features influential in disparate impact based on gender). The associated methodology can be used to drive design of transparency mechanisms as well as internal testing and audit of big data systems.  ***Anupam Datta*** *is an Associate Professor (with tenure) at Carnegie Mellon University where he holds a joint appointment in the Computer Science and Electrical and Computer Engineering Departments. His research area is security and privacy. His current focus is on information accountability -- foundations and tools that can be used to provide oversight of complex information processing ecosystems (including big data systems) to examine whether they respect privacy, and other desirable values in the personal data protection area, such as fairness and transparency. His work has produced accountability tools deployed in industry, and studies that rigorously demonstrate concerns with privacy, fairness, and transparency in online behavioral advertising. He holds a Ph.D. in Computer Science from Stanford University.*
Coffee Break	
10:45–12:25	**Session 1: Privacy engineering tools** **DataTags, Data Handling Policy Spaces and the Tags Language** Michael Bar-Sinai, Latanya Sweeney and Mercè Crosas *Widespread sharing of scientific datasets holds great promise for new scientific discoveries and great risks for personal privacy. Dataset handling policies play the critical role of balancing privacy risks and scientific value. We propose an extensible, formal, theoretical model for dataset handling policies. We define binary operators for policy composition and for comparing policy strictness, such that propositions like "this policy is stricter than that policy" can be formally phrased. Using this model, the policies are described in a machine-executable and human-readable way. We further present the Tags programming language and toolset, created especially for working with the proposed model.* *Tags allows composing interactive, friendly questionnaires which, when given a dataset, can suggest a data handling policy that follows legal and technical guidelines. Currently, creating such a policy is a manual process requiring access to legal and technical experts, which are not always available. We present some of Tags' tools, such as interview systems, visualizers, development environment, and questionnaire inspectors. Finally, we discuss methodologies for questionnaire development. Data for this paper include a questionnaire for suggesting a HIPAA compliant data handling policy, and formal description of the set of data tags proposed by the authors in a recent paper.*

A Semi-Automated Methodology for Extracting access control rules from the European Data Protection Directive

Kaniz Fatema, Christophe Debruyne, Dave Lewis, Declan O'Sullivan, John Morrison and Abdullah Al Mazed

Handling personal data in a legally compliant way is an important factor for ensuring the trustworthiness of a service provider. The EU data protection directive (EU DPD) is built in such a way that the outcomes of rules are subject to explanations, contexts with dependencies, and human interpretation. Therefore, the process of obtaining deterministic and formal rules in policy languages from the EU DPD is difficult to fully automate. To tackle this problem, we demonstrate in this paper the use of a Controlled Natural Language (CNL) to encode the rules of the EU DPD, in a manner that can be automatically converted into the policy languages XACML and PERMIS. We also show that forming machine executable rules automatically from the controlled natural language grammar not only has the benefit of ensuring the correctness of those rules but also has potential of making the overall process more efficient.

Compliance Monitoring of Third-Party Applications in Online Social Networks

Florian Kelbert and Alexander Fromm

With the widespread adoption of Online Social Networks (OSNs), users increasingly also use corresponding third-party applications (TPAs), such as social games and applications for collaboration. To improve their social experience, TPAs access users' personal data via an API provided by the OSN. Applications are then expected to comply with certain security and privacy policies when handling the users' data. However, in practice, they might store, use, and distribute that data in all kinds of unapproved ways. We present an approach that transparently enforces security and privacy policies on TPAs that integrate with OSNs. To this end, we integrate concepts and implementations from the research areas of data usage control and information flow control. We instantiate these results in the context of TPAs in OSNs in order to enforce compliance with security and privacy policies that are provided by the OSN operator. We perform a preliminary evaluation of our approach on the basis of a TPA that integrates with the Facebook API.

Obstacles to Transparency in Privacy Engineering

Kiel Brennan-Marquez and Daniel Susser

Transparency is widely recognized as indispensable to privacy protection. However, producing transparency for end-users is often antithetical to a variety of other technical, business, and regulatory interests. These conflicts create obstacles which stand in the way of developing tools which provide meaningful privacy protections or from having such tools adopted in widespread fashion. In this paper, we develop a "map" of these common obstacles to transparency, in order to assist privacy engineers in successfully navigating them. Furthermore, we argue that some of these obstacles can be successfully avoided by distinguishing between two different conceptions of transparency and considering which is at stake in a given case—transparency as providing users with insight into what information about them is collected and how it is processed (what we call transparency as a "view under-the-hood") and transparency as providing users with facility in navigating the risks and benefits of using particular technologies.

International Workshop on Privacy Engineering

http://ieee-security.org/TC/SPW2016/IWPE

May 26th, 2016 at the Fairmont, San Jose, CA

Co-located with 37th IEEE Symposium on Security and Privacy

12:25-12:30	**Best Paper Award Ceremony**
Lunch	
13:30–14:20	**Session 2: Privacy engineering techniques** **Oblivious Mechanisms in Differential Privacy: Experiments, Conjectures, and Open Questions** Chien-Lun Chen, Ranjan Pal and Leana Golubchik *Differential privacy (DP) is a framework to quantify to what extent individual privacy in a statistical database is preserved while releasing useful aggregate information about the database. In this work, we aim an exploratory study to understand questions related to the optimality of noise generation mechanisms (NGMs) in differential privacy by taking into consideration the (i) query sensitivity, (ii) query side information, and (iii) the presence of longitudinal and collusion attacks. The results/observations from our study serve three important purposes: (i) provide us with conjectures on appropriate (in the sense of privacy-utility tradeoffs) oblivious NGM selection for scalar queries in both non-Bayesian as well as Bayesian user settings, (ii) provide supporting evidence and counterexamples to existing theory results on the optimality of NGMs when they are tested on a relaxed assumption set, and (ii) lead to a string of interesting open questions for the theory community in relation to the design and analysis of provably optimal oblivious differential privacy mechanisms.* **A Critical Analysis of Privacy Design Strategies** Michael Colesky, Jaap-Henk Hoepman and Christiaan Hillen *The upcoming General Data Protection Regulation is quickly becoming of great concern to organizations which process personal data of European citizens. It is however nontrivial to translate these legal requirements into privacy friendly designs. One recently proposed approach to make 'privacy by design' more practical is privacy design strategies. This paper improves the strategy definitions and suggests an additional level of abstraction between strategies and privacy patterns: 'tactics'. We have identified a collection of such tactics based on an extensive literature review, in particular a catalogue of surveyed privacy patterns. We explore the relationships between the concepts we introduce and similar concepts used in software engineering. This paper helps bridge the gap between data protection requirements set out in law, and system development practice.*
14:20–15:15	**Panel: Tools in support of privacy engineering techniques** Anupam Datta, Arvind Narayanan, Sadia Afroz Privacy engineering tools refers to structured (and automated) means to support the use of software engineering methods and techniques to capture and address privacy issues systematically. Tools can be used by software engineers, computer scientists and their teams. They can be of assistance when developing standalone privacy applications (e.g. secure messaging), enhancing privacy of information systems or protocols (e.g., addressing privacy issues in machine learning, internet protocols), or assessing emergent privacy violations that occur in complex environments (e.g., web privacy). In the first panel, we will focus on tools intended to support technical experts in completing specific privacy engineering activities.

	***Anupam Datta** is an Associate Professor (with tenure) at Carnegie Mellon University where he holds a joint appointment in the Computer Science and Electrical and Computer Engineering Departments.* ***Arvind Narayanan** is an Assistant Professor of computer science at Princeton where he researches and teaches information privacy and security, and moonlight in technology policy.* ***Sadia Afroz** is a research scientist at the International Computer Science Institute (ICSI) where she researches security, privacy and machine learning.*
Coffee Break	
15:45–16:50	**Session 3: Privacy engineering methodologies** **Privacy Risk Analysis Based on System Control Structures: Adapting System-Theoretic Process Analysis for Privacy Engineering** Stuart Shapiro *To date, top-down efforts to evolve and structure privacy engineering knowledge have tended to reflect common systems engineering/development life cycle activities. A different approach suggests a particular need for technical analytical methods. To help address this need, this paper proposes to adapt for privacy engineering an existing technique, System-Theoretic Process Analysis (STPA), developed for safety engineering. The foundations of STPA are discussed, its security extension, STPASec, is described, and modifications to STPA-Sec are proposed to produce STPA-Priv. STPA-Priv is then applied to a simple illustrative example.* **Privacy Harm Analysis: A Case Study on Smart Grids** Sourya Joyee De and Daniel Le Métayer *To carry out a true privacy risk analysis and go beyond a traditional security analysis, it is essential to distinguish the notions of feared events and their impacts, called "privacy harms" here, and to establish a link between them. In this paper, we provide a clear relationship among harms, feared events, privacy weaknesses and risk sources and describe their use in the analysis of smart grid systems. This work also lays the foundation for a more systematic and rigorous approach to privacy risk assessment.* **From Privacy Impact Assessment to Social Impact Assessment** Lilian Edwards, Derek Mcauley and Laurence Diver *In order to address the continued decline in consumer trust in all things digital, and specifically the Internet of Things (IoT), we propose a radical overhaul of IoT design processes. Privacy by Design has been proposed as a suitable framework, but we argue the current approach has two failings: it presents too abstract a framework to inform design; and it is often applied after many critical design decisions have been made in defining the business opportunity. To rebuild trust we need the philosophy of Privacy by Design to be transformed into a wider Social Impact Assessment and delivered with practical guidance to be applied at product/service concept stage as well as throughout the system's engineering.*

International Workshop on Privacy Engineering

http://ieee-security.org/TC/SPW2016/IWPE

May 26th, 2016 at the Fairmont, San Jose, CA

Co-located with 37th IEEE Symposium on Security and Privacy

16:50-17:45	**Panel: Tools in support of privacy engineering methodologies** Katie Shilton (Moderator), Aleecia M. McDonald, Sean Brooks, Tony Berman This second panel will present tools that support the execution of methodologies used by teams to manage privacy engineering activities and communicate results to the general public. In the discussion, we hope to reflect on what privacy problem the tools attend to, their potential uses and limitations, as well as forthcoming research challenges related to deployment, extension or evolution of privacy engineering tools. ***Katie Shilton*** *is an Associate Professor at the University of Maryland, leads the Ethics & Values in Design (EViD) Lab at the UMD iSchool, and is the director of the CASCI research center.* ***Aleecia M. McDonald*** *is a privacy researcher and non-resident Fellow with Stanford's Center for Internet & Society where she focuses on the public policy issues of Internet privacy, including user expectations for Do Not Track, behavioral economics and mental models of privacy, and the efficacy of industry self regulation.* ***Sean Brooks*** *represents the Applied Cybersecurity Division of the National Institute of Standards and Technology (NIST) on the west coast. He supports a broad range of NIST projects, including the National Strategy for Trusted Identities in Cyberspace program office and the new privacy engineering program.* ***Tony Berman*** *is a Senior Product Manager at TRUSTe, a leading global Data Privacy Management (DPM) company.*
17:45–18:00	**Wrap-up and concluding remarks**