Report from Dagstuhl Seminar 18181

Towards Accountable Systems

Edited by

David Eyers[1], Christopher Millard[2], Margo Seltzer[3], and Jatinder Singh[4]

1 **University of Otago, NZ,** dme@cs.otago.ac.nz
2 **Queen Mary University of London, GB,** c.millard@qmul.ac.uk
3 **Harvard University – Cambridge, US,** margo@eecs.harvard.edu
4 **University of Cambridge, GB,** js573@cam.ac.uk

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 18181 "Towards Accountable Systems", which took place from April 29th to May 4th, 2018, at Schloss Dagstuhl – Leibniz Center for Informatics. Researchers and practitioners from academia and industry were brought together covering broad fields from computer and information science, public policy and law.

Many risks and opportunities were discussed that relate to the alignment of systems technologies with developing legal and regulatory requirements and evolving user expectations.

This report summarises outcomes of the seminar by highlighting key future research directions and challenges that lie on the path to developing systems that better align with accountability concerns.

Seminar April 29–May 4, 2018 – http://www.dagstuhl.de/18181
2012 ACM Subject Classification Applied computing → Law, Computer systems organization → Cloud computing, Security and privacy, Social and professional topics → Governmental regulations,Social and professional topics → Technology audits
Keywords and phrases accountability, compliance, audit, systems, engineering, cloud computing, internet of things, law, regulation, GDPR, security, privacy, data provenance
Digital Object Identifier 10.4230/DagRep.8.4.126
Edited in cooperation with Jennifer Cobbe

1 Executive Summary

David Eyers (University of Otago, NZ)
Christopher Millard (Queen Mary University of London, GB)
Margo Seltzer (Harvard University – Cambridge, US)
Jatinder Singh (University of Cambridge, GB)

Background and Motivation

Technology is becoming increasingly pervasive, impacting all aspects of everyday life. Our use of apps and online services is tracked and extensively processed (data analytics), and the results are used for various purposes, predominately advertising. Monitoring and surveillance by sensors in smart cities creates vast amounts of data, much of which can be identifiably linked with people. Smart home, health and lifestyle monitoring, and other sensor technologies yield sensitive personal data; mobile phones reveal people's positions, and their calls are

Towards Accountable Systems, *Dagstuhl Reports*, Vol. 8, Issue 04, pp. 126–163
Editors: David Eyers, Christopher Millard, Margo Seltzer, and Jatinder Singh
DAGSTUHL REPORTS Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

tracked leading to data that can be used to determine social linkages and sometimes mental wellbeing. Such collection and analysis of personal data raises serious privacy concerns. A key aspiration is to provide end-users with a means to understand their digital footprints, and control the propagation, aggregation and retention of their data.

Concerns over data movement, location, processing and access have led to increasing regulation, both national and international. An example is the recently adopted EU General Data Protection Regulation (GDPR) that reinforces and expands individual rights, as well as restrictions and obligations regarding personal data. However, data moves easily beyond geographical boundaries, and use of cloud computing resources may mean that stored data may be replicated in multiple locations worldwide, with potential for conflicts between applicable laws and jurisdictions. Governments may demand access to data (whether stored locally or remotely) and this may result in complex legal disputes. Regulations, codes of conduct, and best practices can incentivise the use of particular technical mechanisms for data management. Examples include encryption and anonymisation, for example when using medical data for research. However, there are often misalignments between legal/regulatory aims and the capabilities of the technologies.

Key issues concern how to demonstrate compliance with regulations, such as those regarding how data is handled and used, and, in cases of failure, how to hold the appropriate entities accountable. This is a particular challenge for wide-scale, federated, or cross-border systems. In large or complex systems, data may be handled by many different parties, falling under various management regimes and jurisdictions. Such concerns are not only horizontal (e.g., data being exchanged between parties, across geographic regions) but also vertical, where different levels of the services stack are managed by different parties (e.g., a company application running over a Heroku PaaS that runs over Amazon IaaS). Most end-users (people!) are oblivious to the potential complexity of such systems, let alone the complexity of the legal requirements that underpin such architectures. In general, the lack of transparency and uncertainty about the means for compliance with legal obligations, along with a lack of technical means for managing such concerns, may inhibit innovative technology development (a "chilling factor"), may escalate compliance costs, may trigger inappropriate policy responses, and may work to undermine public trust in technology.

These concerns will only grow in prominence, given the increasing deployment of sensors, generating ever-more data; actuators, giving systems physical effects; and the use of machine learning, facilitating automation. In response, this seminar brought together experts from the computer science and legal communities, spanning academia and industry, to explore issues of accountability as it relates to data and systems. The seminar aimed to: (i) raise awareness of and establish new research directions concerning issues of accountability as they relate to systems, given directions in systems technologies; (ii) explore developing legal and regulatory requirements; and (iii) investigate issues of user empowerment. A key goal was to increase awareness that law, regulation and requirements for data usage, management, security, confidentiality, quality and provenance should align with the technology, and *vice versa*: technologists should be legally-aware and lawyers should be technology-aware.

Seminar Structure

Due to the diverse backgrounds of the participants, the first day was focused on introductions and ensuring that everyone had a common grounding in key topics. This included a series of guided discussion sessions: Lilian Edwards provided an introduction to legal and regulatory considerations, particularly the European Union General Data Protection Regulation (GDPR); Jon Crowcroft introduced emerging technical architectures such as edge computing; Bertram

Ludäscher led a session exploring data provenance; and Ben Wagner introduced broader ethical and social concerns. A motivating case study was also presented highlighting how an apparently enthusiastic view of emerging Internet of Things technologies might obscure a plethora of questionable social and policy implications.

The structure of the week included multiple breakout sessions in which working groups examined particular topics (below) and reported back summaries of their discussions at plenary sessions. The working group sessions were interspersed with an interactive case study session, that focused on the technological compliance concerns of a hypothetical global hotel chain seeking to introduce a series of IoT and cloud technologies in the current regulatory environment, and a session in which participants were able to present their recent research, abstracts for (most of) which are included in this report.

Moving forward

The topics explored by the working groups at the seminar spanned policy, legal and technical considerations. The topics were seeded by the organisers but were ultimately gathered from the participants through a preference allocation process. The chosen topics included:

- *Trust in systems.*
- *Who is, could or should be accountable in complex systems?*
- *Engineering accountable systems.*
- *Is there a place for data provenance in accountable systems?*
- *Anonymity, identity and accountability.*
- *Thinking beyond consent.*
- *Automating the exercising of rights for collective oversight.*

Each group was asked to produce an abstract summarising the key issues, challenges and ways forward from the discussion. These abstracts are included in this report, and indicate many potential opportunities for research.

Generally, it was felt that *this seminar represented only the start of this important discussion.* It is clear that there is a substantial and urgent need for closer interactions between the technical and legal domains, such that (i) the computer science communities better understand the legal requirements and constraints that impact the design, implementation and deployment of technology; and (ii) the legal communities gain more of a grounding in the nature, capabilities, and potential of the technology itself. It was also recognised that there is potential for better collaboration amongst different computer science communities; for example, to have greater interactions between those working in systems, provenance and machine learning.

In light of this, key to moving forward is to work to form collaborative research proposals, and to organise relevant meetings, in order to drive progress on the topics, challenges and research opportunities identified during this seminar. As issues of accountability increase in importance and urgency, it is vital that researchers across academia, industry and civil society work together to proactively confront these challenges.

2 Table of Contents

Executive Summary

Working groups

Overview of Talks

3 Working groups

3.1 Trust in Systems

Jennifer Cobbe (University of Cambridge, GB), Jon Crowcroft (University of Cambridge, GB), David Eyers (University of Otago, NZ), Krishna P. Gummadi (MPI-SWS – Saarbrücken, DE), Joshua A. Kroll (University of California – Berkeley, US), Derek McAuley (University of Nottingham, GB), Michael Veale (University College London, GB), and Michael Winikoff (University of Otago, NZ)

At the outset the group noted the difficulty of defining trust, acknowledging that it means very different things in different contexts. It was noted that this is an inherently interdisciplinary area, and, with participants being from computer science and law backgrounds, it was felt that a proper study of trust in systems would benefit from the input of others from disciplines such as philosophy, psychology, and sociology. It was also noted that there may be circumstances in which distrust (which may not necessarily result in not using a system) is more appropriate than trust. However, trust was generally recognised as being the "beliefs/faith you need to have about the things that you cannot verify", and the group eventually settled on trust in this context as involving "confidence in the behaviour of a system".

Various factors which affect trust/distrust were identified. Social factors may affect which individuals, organisations, and systems are trusted. It was acknowledged that well-developed and properly-enforced regulatory frameworks can help engender trust. People often also rely on mental models and folk theories; this has been shown, for example, in relation to social media feeds, with people becoming upset when their theories fail [1][2]. Contextual factors were also felt to be important, given that the consequences of a system failure can be very different for different people–for example, a robot tripping up an elderly person may be a much bigger problem than if it was someone younger. Cultural factors are also important; concepts of privacy, for example, differ between geographies, communities, and demographics. Finally, it was noted that infrastructure, including security infrastructure, plays a role in engendering trust.

The group moved on to discuss technological features which could be engineered into systems to assist trustworthiness. Predictability, reliability, and repeatability were felt to help manage user expectations of their interactions with systems, as trust may result where systems behave as expected. While transparency was generally felt to be a useful means of improving trustworthiness, it was acknowledged that there are unresolved questions over the extent to which it is useful and over what kind of information should be exposed. Safety marks were considered, but the difficulty of verifying software poses a problem. Engineering for security and privacy were also considered to be important factors in developing trustworthy systems. Finally, giving users control was considered to be an important feature of trust, as it was felt that users may lose trust when a system is beyond their control.

Following this discussion, the case study selected was that of a voice-controlled robotic vacuum cleaner. This interacts in a complex way with its environment, and use of such a device requires trust in the robot, trust in the organisation who makes the robot, and trust in the infrastructure involved in the robot. It was felt important that it would do what was expected, but also that it would not do anything unexpected (it should work when turned on, should stop when it is turned off, etc.). It should take context into account so as to be unobtrusive and safe, could refuse to take actions which would be unsafe, and the consequences of failure should be considered and mitigated so far as possible so as to avoid harm when faults occur. Given that the robot would be voice-controlled, it should be clear about when it is listening and when it isn't, and a clear off switch for the microphone should be available. The organisation who makes the robot should not retrospectively decide to sell user data from devices which were sold on the basis that this would not occur. The group also considered that a button which would explain why the robot had taken an action would help engender trust in its decisions.

In terms of next steps, it was felt that there was a need for significant research in this area, including on what trust means in different contexts, on what factors affect trust and distrust, and on what technical approaches could potentially engender trust so as to help engineer trustworthy systems. It was felt that there was a need for a toolkit which would help people assess their own interpretation of trust and trustworthiness. The group agreed that it was important that trustworthy components or features not be hoarded commercially, to avoid the development of an industry that is paid by the existence of the problem. Finally, the group concluded that it was important to avoid having to repeatedly start this work from scratch–there is a need for work to be done on developing a knowledge base which can be compositional in terms of concerns, intentions, solutions, and so on.

References

1 Motahhare Eslami, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, Alex Kirlik, 2016, *First I "like" it, then I hide it: Folk theories of social feeds*, CHI 2016 – Proceedings, 34th Annual CHI Conference on Human Factors in Computing Systems

2 Taina Bucher, 2017 *The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms*, Information, Communication and Society, 20(1), pp.30–44

3.2 Who is Accountable?

Ben Wagner (Wirtschaftsuniversität Wien, AT), Virgilio Almeida (Federal University of Minas Gerais-Belo Horizonte, BR), Tristan Henderson (University of St Andrews, GB), Heleen Louise Janssen (Ministry of the Interior and Kingdom Relations, NL), Christopher Millard (Queen Mary University of London, GB), and Barbara Staudt Lerner (Mount Holyoke College – South Hadley, US)

Who is or could or should be accountable / responsible in complex systems?

3.2.1 Identify specific issues that more precisely describe it in a specific use case

- **Specific use case**: a connected camera used in a smart home. The camera is used to open the door (for example, to delivery persons) based on facial recognition.

- **Which actors** could or should be accountable for the camera?
  1. Educators of designers and the public
  2. Research ethics boards (in academia and industry)
  3. Standards bodies

4. Regulators, including Data Protection Authorities (DPAs)–acting in determining norms
5. Designers
6. Manufacturers of hardware
7. Developers of software
8. Cloud Service Providers
9. Connectivity providers
10. Social media platforms
11. Installers
12. End-users
13. Regulators, including DPAs–acting in an enforcement capacity
14. Courts

- **Why** are or should these actors be accountable?
  - Knowledge
  - Control
  - Proximity

- **At what stage in the processing** do they become accountable and how can they be made more accountable and responsible?
  - *Ex ante*: before the product design phase starts
  - Production: during the design, development, installation and deployment of the product
  - *Ex post*: after the product has been deployed

These actors are considered accountable because of their knowledge of systems in operation, because of proximity to the system and (at least a certain amount of) control at a certain moment in processing.

3.2.2 Propose mechanisms to address the issue (tech-soc-reg, etc.).

Under what conditions could or should these actors be held accountable and at what stages? To identify mechanisms that could be used to address the issue who is or could or should be accountable / responsible in complex systems, we discussed a variety of different actors in the home camera use case and how they might be made more accountable:

1. The home owner or occupier could be held responsible for ensuring the camera does not cover public spaces.
2. The designer or developer could be held responsible for non-automated opening of the door.
3. The system integrator could be held responsible for operation of the system.
4. The Software as a Service (SaaS) provider could be held responsible for the cloud service (e.g., facial recognition or authentication).
5. The camera designer could be held accountable and incentivised to conduct an ethical impact assessment.
6. The retailer (e.g., a department store such as John Lewis, or even Amazon Marketplace?) could be held accountable for bringing the product to market.

⇒ **All individual solutions can create greater accountability of the system**
⇒ **No one single solution fixes the entire accountability problem**

1. **The home owner or occupier could be made responsible for ensuring the camera does not cover public spaces.**
   - *Ex ante*
     - education, e.g. domestic science teaching in schools could include how to maintain a smart home
   - Production
     - proper installation procedures could prevent the house owner from mis-installing
     - additional information about obligations and potential for misuse provided to home owner (e.g. through the set-up interface or a manual)
     - a big red sticker / smartphone interface informing users that they need to set up the device correctly (e.g., "DO NOT CONNECT THIS DEVICE TO THE INTERNET UNTIL YOU HAVE SET A NEW PASSWORD")
   - *Ex post*
     - liability for privacy violation in a public space (cf. Ryneš case)
2. **Designer or Developer could be made responsible for non-automated opening of the door.**
   - *Ex ante*:
     - engineering training/education to think about challenges of automation
   - Production:
     - information provided in UX during the set-up scheme
     - default set to non-automation
   - *Ex post*:
     - product recall for unfixable product
     - litigation and enforcement action by regulators
3. **System integrator could be made responsible for operation of the system.**
   - *Ex ante*:
     - a clear specification that conforms to all relevant state of art standards (ethical system design, ISO, IEEE)
   - Production:
     - update the system
     - keep users informed about problems/vulnerability and proper usage
     - employ user-centric design and extensive testing of system in controlled environment
     - do not collect data that are not needed (data minimization, select while you collect)
     - obtain user feedback and update product accordingly
   - *Ex post*:
     - professional standards developed further based on reports of system integrator
     - DPO reports of system integrator shared
4. **Software as a Service provider could be made responsible for the cloud service.**
   - *Ex ante*:
     - responsible for security of supply chain
     - cloud standards certification (ISO, CSA)
   - Production:
     - authentication and access controls
     - SaaS providers monitor for security vulnerabilities and notify home owners (push for competitive advantage, market power)
   - *Ex post*:
     - enforcement by regulators and courts

    - breach notification if SaaS leaks data
    - could be forced to do product recall if product cannot be updated or fixed

5. **Camera designer could be accountable and conduct an ethical impact assessment.**
    - *Ex ante*:
        - camera designer follows a relevant code of conduct for software developers
        - designer conducts an ethical impact assessment to assess risks and document them
    - Production:
        - check that they are following the actions that were determined in the impact assessment
        - update impact assessment based on process of development
    - *Ex post*
        - self-motivated reasons to be ethical
        - market pressure to be ethical (e.g. offering a trustworthy/accountable product as a competitive advantage)
6. **Point of sales person (a department store such as John Lewis, or even Amazon Marketplace?) could be accountable for bringing product into market.**
    - *Ex ante*:
        - training for staff
        - sourcing appropriate products (don't sell insecure products)
        - know your supply chain and take responsibility for vendors
    - Production:
        - follow-up information to customer to ensure ethical usage of product and provide support in doing so
        - deploy an ethical chatbot
    - *Ex post*
        - product recall if it becomes insecure and cannot be patched
        - provide lifetime warranty for product and insurance mechanism
        - provide support for customers if they have subsequent problems (e.g. unable to update device)

3.2.3 Evaluate the efficacy of the mechanisms proposed / mapping the complexity:

Potentially contentious issues include:

- Large number of actors–who can be held accountable: designer/toolmakers–people having more knowledge about bottleneck problems?
- Reasonableness of assigning responsibility to tool makers (e.g. maker of motion detector sensor that is incorporated into the camera)
- Dual/multiple use challenge
- Non-legal versus legal mechanisms (What carrots or sticks might be needed to make non-legal mechanisms effective?)
- A conflict between legal norms and ethical frameworks may arise. How paternalistic should systems be to be effective?

3.2.4 Next steps, research directions, gaps in current knowledge

Research directions:

- Are current and proposed impact assessment mechanisms (e.g. GDPR Art 35 DPIA, AINow Algorithmic Impact Assessments) suitable for use by all of the actors in the system?
- Could there be certifications associated with the technology, and what organisation would do those certifications? For example, certification assuring that all data are encrypted.
- It might be beneficial to have standard APIs to support integration of IoT devices. For example, how to control a lock. Who should define those? Can there be some certification about whether a device satisfies those interfaces?
- What makes an actor accountable, e.g. are knowledge, proximity and control the most appropriate benchmarks to determine accountability? Can accountability vary/change in the process of operation? How does this relate to the role of the data controller in the GDPR, and to legal liability?

References

1 Opinion 3/2010 on the principle of accountability, Data Protection Working Party Article 29 (adopted on 13 July 2010)

2 Opinion 1/2010 on the concepts of "controller" and "processor", Data Protection Working Party Article 29 (adopted on 16 February 2010)

3 A. Crabtree, T. Lodge, J. Colley, C. Greenhalgh, K. Glover, H. Haddadi, Y. Amar, R, Mortier, J. Moore, L. Wang, P. Yadav, J. Zhao, A. Browns, L. Urquharts and D. McAuley. Building Accountability into the Internet of Things: The IoT Data Box Model, in: *Journal of Reliable Intelligent Environments* (2018) 4, p. 39–55. See https://link.springer.com/content/pdf/10.1007%2Fs40860-018-0054-5.pdf.

4 R.H. Weber, Accountability in the Internet of Things, in: *Computer Law and Security Review* (2011), Volume 27, Issue 2, April 2011, Pages 133–138. https://doi.org/10.1016/j.clsr.2011.01.005.

5 L. Urquhart, T. Lodge and A. Crabtree, *Demonstrably doing accountability in the Internet of Things*, School of Computer Science, University of Nottingham, UK. See https://arxiv.org/pdf/1801.07168.pdf.

6 W. Kuan Hon, E. Kosta, C. Millard and D. Stefanatou, *Cloud Accountability: The Likely Impact of the Proposed EU Data Protection Regulation*, Queen Mary School of Law Legal Studies Research Paper No. 172/2014; Tilburg Law School Research Paper No 07/2014. See https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2405971

7 W. Kuan Hon, C. Millard and J. Singh, *Twenty Legal Considerations for Clouds of Things* (January 4, 2016). Queen Mary School of Law Legal Studies Research Paper No. 216/2016. Available at SSRN: https://ssrn.com/abstract=2716966 or http://dx.doi.org/10.2139/ssrn.2716966.

3.3 Engineering Accountable Systems

Martin Henze (RWTH Aachen, DE), Virgilio Almeida (Federal University of Minas Gerais-Belo Horizonte, BR), Jean Bacon (University of Cambridge, GB), Melanie Herschel (Universität Stuttgart, DE), Maximilian Ott (CSIRO – Alexandria, AU), Frank Pallas (TU Berlin, DE), Thomas Pasquier (University of Cambridge, GB), Silvia Puglisi (The Tor Project – Seattle, US), Margo Seltzer (Harvard University – Cambridge, US), Michael Winikoff (University of Otago, NZ), and Martina Zitterbart (KIT – Karlsruher Institut für Technologie, DE)

Also from an engineering perspective, accountable systems raise certain challenges and research questions. In particular, we feel that engineers who should implement accountable systems require support in doing so. The more engineering-focused aspects of accountable systems were discussed in two consecutive working groups which are jointly reported on here. In both groups, the discussions on how to engineer (complex) accountable systems were guided by the question of what are the best practices, design guidelines, implementation guidelines, operational guidelines, and re-usable technical primitives required to provide transparency of operation, allow for verification against a specification, and assign responsibility for actions. We subsume respective research and development activities under the term "Accountability Engineering".

In particular, we foresee the engineering of accountable systems to become relevant in settings comprising multiple actors and stakeholders with "accountability chains" having to be established across system, organizational, domain, and even regulatory boundaries. For instance, we envision scenarios from the field of logistics, where multiple companies are involved in the distribution of goods–employing technologies like autonomous and connected vehicles–and where responsibility must be assigned to the correct party as soon as shipped goods turn out to have been damaged (broken, not sufficiently cooled, ...) somewhere during the transportation or supply chain.

When discussing accountability in such scenarios, it is important to keep in mind that there are different perspectives and hence different motivations to cater for accountability. Motivations for striving towards accountability vary significantly–ranging from regulatory compliance to business-driven risk-benefit weighings (e.g., when deciding about whether to use a certain service or not [1]) and ethical considerations. Independently from this, however, an area that we believe has not received appropriate consideration so far is considering accountability as a higher level service or system property/quality for which one can charge extra or which can be used to differentiate from competitors in competitive markets.

As a foundation for realizing accountable systems using Accountability Engineering, we require a common understanding of "accountable systems" specifically for engineers. Here, we particularly distinguish between three different aspects of accountability which must be reflected in Accountability Engineering:

- Transparency: Accountable systems must provide evidence about relevant facts concerning involved systems [2, 3] as well as external events and incidents. Systems themselves as well as provided evidence must be auditable and must facilitate justifiability to multiple parties (and not only to the one party operating a system).
- Verifiability: Beyond mere transparency, accountability also requires that stakeholders are able to verify aspects such as the compliance with regulatory givens, the conformance

with agreements, etc. Accountable systems must reflect this need in the collection and provision of evidence.

- Responsibility: Finally, the concept of accountability is closely related to responsibility [4]. Aspects such as assigning 'blame'? to the party actually responsible for a damage or the enforcement of reparations must be technically supported by accountable systems. Last but not least, this also includes the need for being able to 'fix' systems after a damage or malfunction has actually occurred.

To appropriately address all these aspects, Accountability Engineering requires a broad, multidisciplinary understanding that combines knowledge from domains such as technology, business, law, and others. In many cases, this understanding will also have to cover multiple contextual environments (e.g., when an international delivery or supply chain comprises multiple legal/regulatory regimes). As a further precondition for making Accountability Engineering successful, it is important that engineers understand that accountability is an integral part of a product or service specification and not an add-on that can be (easily) added at a later point of time, possibly even at an extra charge.

On this basis, we envision the field of Accountability Engineering to be concerned with important research questions such as

- What information to provide to an engineering team that is about to embark on an accountable systems project?
- What to teach students about how to build accountable systems?
- What regulations are in place that have implications for the system that is to be built?
- How to incorporate regulatory changes into the life-cycle of a product or service?
- How to manage different context-specific regulatory regimes?
- How to consider ethics when designing, implementing, and operating accountable systems?

Even though considerably related, Accountability Engineering will differ from traditional software or systems engineering in multiple respects. In particular, it will be shaped by a strong focus on multidisciplinary stakeholder analysis (calling for approaches such as i* [5] or GRL [6], which capture stakeholders, their goals, and their dependencies on each other, and on the system-to-be).

Furthermore, the collection, flow, and use of evidence data as well as questions around proving the trustworthiness and provenance of such data will play a dominant role [7]. Different from other fields, the disclosure of such data across organizational boundaries will foreseeably be essential in the context of Accountability Engineering. This will lead to many challenges regarding organizations' unwillingness to share internal data with business partners or–e.g., in case of a dispute going to court–with third parties [1]. However, achieving accountability and, thus, desirable overall outcomes, necessarily requires some evidence information to be revealed in a trustworthy and non-disputable manner. Especially for contexts involving multiple cooperating parties, there will also be a need for technically implemented 'evidence chains' (e.g. provenance record [8], information flow audit [9]) that make respective data available to multiple stakeholders in different granularities–e.g., allowing the final recipient of a perishable good to access all temperature measurements that indicate conditions that it was exposed to during transit while another party only sees values exceeding a certain, predefined temperature corridor.

Meeting such requirements in concrete technical systems will call for novel technical building blocks beyond measures already established in other contexts (such as security, for example) [10]. In particular, we consider further research on the following technical mechanisms to be highly important and valuable for the establishment of accountable systems and, thus, to play an integral role in the field of Accountability Engineering:

- Access control for different actors at different granularities: To achieve technically implemented accountability across organizational boundaries, one party (A) must be able to access (some) data collected by and possibly internal to another party (B). At the same time, it must be prevented that information about third parties (e.g., B's customers) are leaked through respective mechanisms.
- Demonstrably trustworthy data capture, storage, and aggregation: Parties will have to be able to demonstrate that they took reasonable effort in making their systems "correct", secure, and tamper-"proof". Noteworthily, "correctness" and "proof" are not necessarily to be understood in the rather strict meaning established in computer science. Instead, Accountability Engineering will presumably also refer to and significantly profit from the more differentiated understanding established in the legal domain.
- Verifiable (possibly distributed) ledgers at the point of interaction between different parties: Whenever accountability-related data crosses organizational boundaries in an evidence chain, there is the risk of data being manipulated in the interest of downstream parties, especially when an incident actually happened. Distributed ledger technologies are a promising approach for ensuring unalteredness of evidence data in such settings and could provide several benefits over traditional public key infrastructure schemes–especially in settings involving a multitude of potentially mutually mistrusting parties.
- Mechanisms for verifying / assessing the "truthfulness" of data in chains of evidence: Being able to assess the truthfulness of evidence data provided throughout evidence chains is essential for achieving accountable systems. Besides technologies that have long been discussed in the context of Trusted Computing, other mechanisms such as advanced plausibility checks will foreseeably also play a significant role in this regard. To be practically relevant, any such mechanism must also support different granularity layers so that necessary abstractions or data transformations (see first point above) do not completely hinder truthfulness assessments.
- Mechanisms and approaches for representing and interacting with cross-organizational chains of evidence on different levels of detail: Besides assessing truthfulness, chains of evidence must also be represented and interacted with for multiple purposes to achieve accountability. In particular, this includes identifying the causes of damage that may be recognized at later stages, demonstrating compliance with regulatory requirements or business agreements, etc. Such interactions must be possible at different levels of detail for technical and non-technical users as well as by automated means. This necessitates a broad range of technical mechanisms for querying and exploring evidence data that are particularly tailored to accountability-related problems.

On this more technology-focused level, we therefore envision Accountability Engineering to deal with additional questions, including:

- What specifically needs to change in the development process to realize accountable systems?
- What type of accountability queries will accountable systems (typically) have to answer?
- Does it suffice to capture (in machine processable form) only "base-level" facts, or are higher-level facts such as commitments between parties, contracts, etc. also necessary?
- What impact can be expected from the existence of autonomous systems? Do they require paradigmatically different information to be captured and provided? Is there a need for "explanation" of "decisions" autonomously being made by such systems?
- Assuming adequate evidentiary capture, how do we support accountability queries, including verification as well as the various forms of exploration and navigation outlined above?

These and many further engineering-related questions will have to be answered on the way towards actually accountable systems. As outlined above, however, the field of Accountability Engineering is inherently transdisciplinary and cannot be mastered from the perspective of engineering (technical systems) alone. Of the many possible paths coming into consideration as next steps towards the establishment of a novel field of Accountability Engineering, we therefore want to particularly highlight the following two:

- Conduct case studies together with experienced lawyers in the context of specific scenarios, rather than abstract considerations: In the course of such case studies, accountability-related conflicts should be anticipated as going to court. Based on a set of assumptions defined to avoid catch-all "It depends" conclusions, respective disputes should then be "emulated" as running through usual legal processes and procedures. On this basis, it should then be explored how specific technologies could change things to the better or even to the worse.
- Establish a research community exploring the increasingly relevant field of Accountability Engineering from different angles, including–at least–technical primitives, specific implications for engineering processes, legal requirements and implications, and educational aspects: In the light of past experiences with transdisciplinary research on novel technologies, we foresee a multitude of challenges to be overcome in this regard: Fostering participants' willingness to actually accept and get into the possible contributions of "other" disciplines; avoiding the assumption of unchangeable or unrealistically overestimated givens and requirements from the "own" discipline; counteracting ever-repeating paths of argumentation without progress; establishing a critical mass of community members sufficiently experienced in multiple domains and so forth. All these challenges notwithstanding, fostering transdisciplinary activities is indispensable for paving the way towards accountable systems that actually matter in practice.

References

1 Frank Pallas. *An Agency Perspective to Cloud Computing*. In: 2014 International Conference on Grid Economics and Business Models (GECON), (pp. 36–51). Springer, 2014

2 Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson, and Harlan Yu. *Accountable Algorithms*. 165 U. Pa. L. Rev. 633 (2016–2017)

3 Martin Henze, Daniel Kerpen, Jens Hiller, Michael Eggert, David Hellmanns, Erik Mühmer, Oussama Renuli, Henning Maier, Christian Stüble, Roger Häußling, and Klaus Wehrle. *Towards Transparent Information on Individual Cloud Service Usage*. In 2016 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), (pp. 366–370). IEEE, 2016.

4 Derick W. Brinkerhoff. *Accountability and health systems: toward conceptual clarity and policy relevance*. Health policy and planning 19.6 (2004): 371–379.

5 *Early requirements: i** [http://istar.rwth-aachen.de/tiki-view_articles.php]

6 *GRL* [http://jucmnav.softwareengineering.ca/foswiki/UCM/DraftZ151Standard]

7 Andreas Haeberlen. *A Case for the Accountable Cloud*. ACM SIGOPS Operating Systems Review, 44(2):52–57, 2010.

8 Thomas Pasquier, Jatinder Singh, Julia Powles, David Eyers, Margo Seltzer, and Jean Bacon. *Data provenance to audit compliance with privacy policy in the Internet of Things*. Personal and Ubiquitous Computing 22, no. 2 (2018): 333–344.

9 Thomas Pasquier, Jatinder Singh, Jean Bacon and David Eyers. *Information flow audit for PaaS clouds*. In IEEE International Conference on Cloud Engineering (IC2E), pp. 42–51, 2016.

10 Denis Butin, Marcos Chicote, and Daniel Le Métayer. *Strong Accountability: Beyond Vague Promises*. Reloading Data Protection: Multidisciplinary Insights and Contemporary Challenges, Springer, pp.343–369, 2014.

3.4 Provenance for Accountable Systems

Lilian Edwards (The University of Strathclyde – Glasgow, GB), Melanie Herschel (Universität Stuttgart, DE), Bertram Ludäscher (University of Illinois at Urbana-Champaign, US), Ken Moody (University of Cambridge, GB), Thomas Pasquier (University of Cambridge, GB), and Jatinder Singh (University of Cambridge, GB)

Provenance generally refers to metadata that describes the production process of some end product. The W3C PROV standard [3] defines provenance as "a record that describes the people, institutions, entities, and activities involved in producing, influencing, or delivering a piece of data or a thing." [2, 1]

In the context of the accountability of systems that involve information technology, we are particularly interested in the provenance of outcomes of (partly) digital processing, e.g., decisions based on machine learning, *ad hoc* interactions in smart environments, output data of big data processing pipelines, or actuations in cyber-physical systems. This provenance includes for instance the processing history, starting from input data via intermediate results, all the way to the final outcome. Indeed, provenance may offer a useful source of audit and 'evidence' that has the potential to associate different actors of a process to their actions, indicate and possibly verify alignment of the actual processing with expectations, and assist compliance with legal obligations [4, 6]. More generally, provenance may support us in achieving transparency, verifiability, and indicating those responsible (or at least, where further investigation or explanations are due), which are three important dimensions of accountability. Provenance may be important to establishing strict (i.e., non-fault-based) legal liability: see the EC Product Liability Directive article 6 which states directly that "a product is defective when it does not provide the safety which a person is entitled to expect" and that this depends in part on "the use to which it could reasonably be expected that the product would be put" and the time when the product was put into circulation–both of these may depend on recording provenance. Based on an illustrative use-case, we argue that provenance plays a critical role in raising levels of accountability in systems. However, technical and legal challenges remain, and the interplays between the technical and legal aspects require further consideration.

In the following, we focus on the technical challenges, discussed based on an example use case.

3.4.1 Use Case: Ambient Assisted Living

As an example illustrating the potential benefits of provenance for accountable systems, let us consider a smart-home environment targeted towards ambient assisted living. This environment aims at supporting elderly people or people with disabilities in maintaining their independence by living at home as long as possible. Through the analysis of data collected by sensors, video, and context information provided for example through a medical history, and

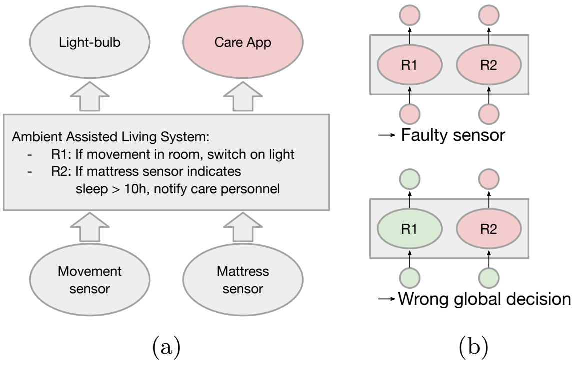
Light-bulb
Care App
Ambient Assisted Living System:
- R1: If movement in room, switch on light
- R2: If mattress sensor indicates sleep > 10h, notify care personnel
Movement sensor
Mattress sensor
(a)
R1
R2
→ Faulty sensor
R1
R2
→ Wrong global decision
(b)

Figure 1 Examples of provenance in a sample ambient assisted living use case.

artificial intelligence, supportive actions may be automated. For instance, switching off the stove when nothing is cooked, setting off an alarm when a person remains suspiciously long in bed, switching on dimmed light when movement is noticed on the way to the bathroom at night, sending notifications to care givers, coordinating medical care, etc. Essentially, as depicted in Figure 1, data is collected and transmitted to the analysis system, which then triggers actions on various devices (e.g., the stove, the light bulb, or a smartphone app receiving notifications).

Given this scenario, imagine that at some point, a family member notices that no notification was sent, even though the person was spending too much time in bed. How can this system behavior be accounted for? Clearly, there are many possible causes to this (faulty) system behavior: sensors may be faulty, the data from sensors was not processed, contradictory data was processed leading to an uncaught exception in processing, the analysis system failed, etc. With the availability of proper provenance traces that track any data processing in the ambient assisted living environment, we can actually narrow down the causes for the system behavior. Different provenance scenarios are depicted in Figure 2, where green coloring indicates that data was traced in a component and red indicates that no data was associated with a component. In the upper figure, we see that no data is available at the sensor level, clearly indicating a faulty sensor. The lower figure shows that sensors produced data that was processed but then translated to another (possibly wrong action). This indicates that the decision model made a wrong decision.

Given the illustrative example above, we see that provenance indeed is important metadata for accountability. But how to design systems that collect such provenance? In the following, we propose three system architectures that may be used to that effect.

3.4.2 Architectures for Provenance Capture

To model systems such as the one described above, we consider sensors and smart devices, the decision making system, and the vendor of the smart-home solution. A first solution to capture provenance would be a closed system, where the (proprietary) devices communicate directly with the vendor's cloud where the decision making, and the provenance capture, is based. Clearly, this offers no or very little visibility of the data and provenance, and raises privacy and security concerns. A second architecture involves a personal hub located in the home network that mediates between the smart-devices and the vendor cloud and that performs the decision making, thus putting the computing closer to the edge [5]. This way, only information that is relevant to the vendor services may be communicated, improving the privacy compared to the first solution. Finally, a third option is to extend the capabilities of smart devices themselves, leveraging their increasing computing capabilities to perform computations right at the source and support provenance capture and communication. The three architectures have different characteristics concerning privacy, transparency, control of the data, disclosing different information available to capture provenance.

3.4.3 (Provenance) Data Disclosure

More generally, the information available to generate provenance data will be dependent on the will and incentives for parties to collaborate in transparently disclosing their actions. An ideal situation would be fully transparent parties willing to disclose any details of data processing or decision making. At the opposite end of the spectrum are uncollaborative parties refusing to participate in the generation of provenance data. In such a scenario provenance can only be inferred from the observation of systems events (or records thereof) that are within reach of the users (e.g., recording network activity). Neither end of the spectrum is desirable: total transparency presents security, privacy and business risk, while a total lack of transparency may reflect badly on the concerned party and a certain level of transparency is legally required. It remains to be determined what level of detail is required to permit the identification of a problem's root cause, while limiting risk where privacy, security and competitive advantage are concerned. We may even go further so as to state that one may want fine-grained provenance for internal root cause identification, while disclosing a more abstract representation to other parties.

3.4.4 Summary and Challenges

The above considerations demonstrate that provenance is valuable information for accountable systems, e.g., as a source of evidence that can assist in indicating responsibility, or help in debugging and fixing a system. However, various challenges lie ahead on the path to practically collecting and using provenance in accountable systems. Interesting research questions include: Can provenance be privacy preserving? How can we meet requirements for accountable systems without disclosing additional information? How can we guarantee the availability of the provenance data? What are the legal requirements regarding the management of provenance data? Similarly, how can we ensure the integrity of collected provenance? Also, to capture and use provenance in a complex system involving various parties, how can we support interoperability (at all levels: system, syntactic, semantic)?

References

1 Lucian Carata, Sherif Akoush, Nikilesh Balakrishnan, Thomas Bytheway, Ripduman Sohan, Margo Seltzer, and Andy Hopper. A Primer on Provenance. *Communications of the ACM*, 57 (5): 52–60 (2014)

2 Melanie Herschel, Ralf Diestelkämper, and Houssem Ben Lahmar. A survey on provenance: What for? What form? What from?. *The VLDB Journal*, 26 (6): 991–906 (2017)

3 Luc Moreau and Paolo Missier, eds. PROV-DM: The PROV Data Model. *W3C Recommendation. https://www.w3.org/TR/prov-dm*, 2013

4 Thomas Pasquier, Jatinder Singh, Julia Powles, David Eyers, Margo Seltzer, and Jean Bacon. Data provenance to audit compliance with privacy policy in the Internet of Things. *Personal and Ubiquitous Computing*, 22 (2): 333–344 (2018)

5 Weisong Shi, Jie Cao, Quan Zhang, Youhuizi Li, and Lanyu Xu. Edge computing: Vision and challenges. IEEE Internet of Things Journal, 3(5): 637–646 (2016)

6 Jatinder Singh, Jennifer Cobbe, and Chris Norval. Decision Provenance: Capturing data flow for accountable systems. *arXiv preprint arXiv:1804.05741, https://arxiv.org/abs/1804.05741*, (2018)

3.5 Anonymity, Identity and Accountability

Jean Bacon (University of Cambridge, GB), Heleen Louise Janssen (Ministry of the Interior and Kingdom Relations, NL), Joshua A. Kroll (University of California – Berkeley, US), Silvia Puglisi (The Tor Project – Seattle, US), Jatinder Singh (University of Cambridge, GB), and Martina Zitterbart (KIT – Karlsruher Institut für Technologie, DE)

Group questions: *Does accountability necessitate the identification and/or visibility of the actors involved?* and *In what contexts can a violation of privacy be justified to achieve accountability?*

3.5.1 Considerations about the relation between accountability and identity

Identity has a clear relation to accountability. Identity is important as a means for determining those to hold to account, which can be challenging in a complex system. Does accountability necessarily mean that it is associated with an identifiable person? Identification also poses considerations with regards to the 'users' of a system. Does an individual need to be identifiable–explicitly, or through a pseudonym? Or can (or should?) systems be built in a way so that the details of the individual's identity are not required while still being accountable? And how do the design decisions relating to identity impact on the levels of accountability that systems help to support. Ultimately, 'it depends', issues of identity and anonymity relate to the particular context.

Considerations on accountability: If a system is to support accountability, it needs to record sufficient data to identify relevant events and actors as required by its specification. In a system-of-systems, this accountability data may be partitioned according to the system components and their functionalities. We assume that access to audit data is controlled so that only authorised parties can see it. We assume that a case has been made and authorised for the data to be gathered for legitimate purposes. In this section we assume, at least initially, that the audit data are correct.

3.5.2 Use case: check-in/check-out in public transportation systems

The group identified a use case in which the issues that come with accountability, anonymity and identity can be concretely explored. To investigate options available and trade-offs that might be made, it considered an urban transport system such as the Paris Metro or the Boston T. The group supposed that the initial system design allows people to be completely anonymous, for example, if they choose to use cash to buy tokens. A token will buy any length of journey at any time, and deposited on entry–no other information is available. Suppose such a system is due to be upgraded to meet the following requirements:

- Start and end points of journeys are to be recorded to aid resource planning.
- Different charging models for different journey lengths, different times of day, days of the week, etc.

Consequences of these changes (as well as aiding modelling and planning, as above) are that:

- Individual actions can be inferred–their home and work locations and travel habits.
- Police can be assisted in identifying who could have been present when some crime took place at some location.

If a system is to support accountability it needs to record sufficient data to identify relevant events and actors as required by its accountability specification. In complex system environments this accountability data may be partitioned according to the system components and their functionalities. We assume that access to audit data is controlled so that only authorised parties can see it. We assume that a case has been made and authorised for the data to be gathered for legitimate purposes. In this section we assume, at least initially, that the audit data is correct.

Tensions between accountability on one hand and privacy on the other may arise in these systems, as well as the question of how these tensions can be managed. Specifically *how can accountability in check-in/check-out in city public transportation systems be managed?* Three main systems were considered:

- Token-systems (system as used in the Paris metro). With a one-off deposit of a generic token, people travel unidentifiably.
- Card based systems (London Oyster Card system or credit card system). People can be identified with credit cards and whenever the London Oyster Card is connected to an identifiable person.
- Biometric system (e.g. a future airport with seamless flow systems), where people are identifiable, e.g. by cameras, device traces, other sensors, etc. This use case was not discussed in detail; the group took note of its emergence in the near future.

3.5.3 Drivers for the use of check-in&out systems

Transport related: The group discussed which specific drivers for 'change' exist –e.g. for taking the step away from token-based systems to those entailing check-in/check-out. Taking the Paris Metro with anonymous paper tickets as a starting example, the group noted that original purpose–solely securing that travellers pay for their journey–was less amenable towards city goals of better transport planning and management. As such, many public transport systems use check-in/check-out ticketing approaches: journey tracking provides for insight in understanding travelling patterns, potentially leading to saving money and better allocation of resources (e.g. by offering different services during rush hours for higher prices). The question arose to what extent a traveller's identity is required for such insights. The group questioned whether the use of identities is necessary for transport planning. There may be different levels of aggregation regarding the identities used, or different methods for tracking flow of passengers once travelling, potentially leading to less of an impact on privacy/anonymity. For the detection of fraud by the transport organisation, knowledge of the identity of the travellers *might* assist–depending on the specifics of its implementation.

Not transport related: For crime prevention, detection and prosecution, knowledge of someone's identity may be helpful as well, but it is not clear to what extent an identifiable ticket necessarily assists investigation. Generally, data mining and analytics, to give more information of the flow of people throughout the city (beyond transport purposes) was also identified as a driver.

Other drivers for check-in&check-out are that alternative forms of ticketing can be used, and services can be built on top. For instance, using personal credit cards for ticketing could (by tracking travellers), mean additional services may be offered, such as (targeted) services such as advertisements for trips, experiences, food, etc. In order to realise these customised services, some degree of traveller identity will be required. At present, it appears that privacy and anonymity are being eroded due to commercial and financial incentives.

The group concluded that any next step seems at least to result in increased tracking possibilities and therefore a greater impact on privacy and anonymity.

Typical issues: The group identified typical issues and concerns coming with the use of identifiable check-in/check-out in city public transportation systems:

- All methods can bring statistics
- Public acceptability (agency, personal autonomy, self-determination)
- Identifiability/re-identifiability
- Simplified analytics
- Facilitated 'function creep'
- Discouraging bad-behaviour (gaming the system)
- Anonymous travel will become less self-evident
- Unclear what more exactly happens to the data
- (Prior) design decisions (and constraints) are important
- Fundamental change of a system once in operation is very expensive

3.5.4 Mechanisms to address the tension between accountability and privacy in transportation systems

The main question considered was: *Does accountability necessitate for identification and/or visibility of travellers for the actors?* The group discussed the token and card-types of check-in and check-out systems in public transportation, and how they could be made accountable while respecting the privacy (anonymity and identity) of the individual. The group identified various types of cards, including:

- Fully-identified fare cards
  - Contactless credit card payments, ID cards (based on biometrics)
- Pseudonymous fare cards
  - e.g. Oyster cards
- Fully anonymous electronic fare cards
  - anonymous card paid by e-cash/Zcash
  - payments, trustworthiness of anonymity

A key takeaway was that while all solutions bring statistics, no one solution deals with all accountability and privacy concerns. In short, *there is a trade-off.* Potential barriers to improvements and change in city public transportation systems include:

- Fundamental change of a system once in operation is very expensive
- The increasing number of actors
- Challenges of dual/multiple use (i.e. data and services beyond transport)

3.5.5 General takeaway: context is key

The group concluded that context is key, highlighting some general questions and considerations regarding the identity–anonymity–accountability tradeoffs:

- Determine the actors involved in the processing of the personal data
- What values are in play? How much protection should each value deserve?
- Discrimination (e.g. when means are unreasonably allocated to certain zones based on data, or when certain passengers are unreasonably targeted/checked)
- Agency/personal autonomy of the individual

- Broader benefits of individual-oriented analytics
- The necessity of individual-oriented analytics, are aggregates as effective?
- Security implications, both technical and societal
- Preventing bad behaviour, defining the threat model
- What is necessary to record (considering each actor, each purpose)?
- How to resolve conflicts regarding data retention and data protection?
- How can technology contribute to support oversight practices while respecting privacy/anonymity?
- ID-management: what are the technical, management and 'user' considerations
- Reliability and accuracy of identification mechanisms

References

1 Pfitzmann, Andreas, and Marit Hansen. "A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management." (2010).

2 Branscomb, Anne Wells. "Anonymity, autonomy, and accountability: Challenges to the first amendment in cyberspaces." *The Yale Law Journal* 104.7 (1995): 1639–1679.

3 Christopherson, Kimberly M. "The positive and negative implications of anonymity in Internet social interactions: "On the Internet, Nobody Knows You're a Dog"." *Computers in Human Behavior* 23.6 (2007): 3038–3056.

3.6 Post-Consent

Jennifer Cobbe (University of Cambridge, GB), Martin Henze (RWTH Aachen, DE), Bertram Ludäscher (University of Illinois at Urbana-Champaign, US), Ken Moody (University of Cambridge, GB), Maximilian Ott (CSIRO – Alexandria, AU), and Margo Seltzer (Harvard University – Cambridge, US)

The group agreed that there seems to be a trend in academia, industry, and government to frame issues around the appropriate use of personal data in terms of consent, often drawing on experiences of the medical domain. However, the digital domain has characteristics which make it quite different from the medical domain; sufficiently so that it seems prudent to question the applicability of consent as a basis for processing personal data. Indeed, a significant body of research indicates that relying on consent for the processing of personal data is flawed in many ways [8, 7, 10, 11, 2, 9, 6, 4].

Unlike in the medical domain, users in the digital realm are subjected to a constant onslaught of consent requests from web sites and apps, each with their own privacy policy, as opposed to the comparatively small number of such documents presented by a trusted medical professional. Privacy notices are often very long, confusing, legalistic, and use euphemisms to describe the processing in question (e.g. describing the sale of user data as "sharing"). Given the length, complexity, and number of privacy policies it is unlikely indeed that anybody reads all of them. Indeed, a 2008 study concluded that it would take ten full days to read all the privacy policies with which users are confronted each year [5] (that number presumably having grown greatly in the last decade). Not only is it clear from the literature that people usually do not read privacy policies and often do not understand what they are consenting to, but research has also shown that whether users give consent is heavily dependent on contextual clues [1] (giving data controllers significant influence over whether data subjects give consent to processing). The use of manipulative or deceptive practices in gaining the consent of data subjects is well-documented.[1] Finally, a problem which arises in the digital domain which does not necessarily arise in the medical domain (unless machine learning tools are being used), is that personal data are frequently used to build predictive models, the outputs of which are unforeseeable for data subjects and which may cause their own kind of privacy harms [3]. Given this, it is not clear how one could adequately inform and seek consent from data subjects for such processing.

As a result of these various issues, the group felt that, in order to develop accountable systems of the future, better approaches to providing a lawful basis for processing personal data which respects and protects the rights to privacy and data protection of data subjects is needed. Identifying potential legal bases which provide meaningful control for data subjects as well as developing technical means to protect personal data but which do not rely on consent were thought to be key areas for further research.

References

1 Acquisti, A., Brandimarte, L, and Lowenstein, G., 2015, *Privacy and human behaviour in the age of information*, Science, 347(6221), 30 January, pp.509–514

2 Brandimarte, L., Acquisti, A., Loewenstein, G., 2010, *Misplaced confidences: privacy and the control paradox*, Ninth Annual Workshop on the Economics of Information Security

3 Crawford, K. and Schultz, M., 2013, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, New York University Public Law and Legal Theory Working Papers, Paper 429.

4 Doteveryone, 2018, *People, Power and Technology: The 2018 Digital Understanding Report*

5 McDonald, A. M. and Cranor, L. F., 2008, *The Cost of Reading Privacy Policies*, I/S: A Journal of Law and Policy for the Information Society, 4(3), pp.541–565

6 Mantalero, A., 2015, *The future of consumer data protection in the E.U. Rethinking the "notice and consent" paradigm in the new era of predictive analytics*, CLSR, Vol. 30, pp.643–660

7 Moores, T., 2005, *Do consumers understand the role of privacy seals in e-commerce?*, Communication of the ACM, 48(3), pp.86–91

8 Privacy Leadership Initiative, 2001, *Privacy Notices Research Final Results*, Harris Interactive, 2001

9 Solove, D., 2013, *Privacy Self-Management and the Consent Dilemma*, 126 Harvard Law Review, p.1880

10 TRUSTe, 2006, *Consumers have a false sense of security about online privacy–Actions inconsistent with Attitudes*, TNS, 2006

11 Turow, J., Hoofnagle, C.J., Mulligan, D.K., Good, N., 2007, *The federal trade commission and consumer privacy in the coming decade*, ISJLP, No. 3, pp.723–749

[1] See, for example, https://darkpatterns.org/

3.7 Automating Data Rights

Michael Veale (University College London, GB), Lilian Edwards (The University of Strathclyde – Glasgow, GB), David Eyers (University of Otago, NZ), Tristan Henderson (University of St Andrews, GB), Christopher Millard (Queen Mary University of London, GB), and Barbara Staudt Lerner (Mount Holyoke College – South Hadley, US)

This working group examined how individual rights in the law might, through computationally-mediated means, operate automatically and at scale, and the effect this might have on accountable data-driven systems more generally. The report that follows should be read as as a discussion piece rather than a finalised piece of work.

3.7.1 Discussed Problems

Today, a wide array of entities hold, transform and share personal data. The networks established through these processes are complex, and present a major barrier to accountable systems [4, 14]. Not only do individuals rarely know what data are held by organisations and the inferences made about them, but oversight bodies and regulators often lack the expertise to assess processing practices. As a result, it is likely that many of these actors are operating outside both the letter of the law and are violating consumers' reasonable expectations. This is likely to continue to be the case, even notwithstanding the higher stakes under the General Data Protection Regulation (GDPR)[2] of fines of €20m / up to 4% of global turnover.

The core problem is how and whether, without significant legal change, the existing provisions might better serve consumers in fostering accountability within these systems.

3.7.2 Augmenting Individual Rights

Data protection rights and obligations are powerful in theory, providing building blocks for control of a vast array of data types and processing practices relating to an individual. Many powerful rights are present in current European data protection law. The right to access data that relate to an individual, and the right to "portability" of a narrower range of personal data, both enable a data subject to take copies of data and use them for their own purposes. The right to erasure, often referred to as the "right to be forgotten", is a qualified right to ask a controller to delete or obscure data (for example, by delisting specific search engine results). There are a variety of ways that a data subject can prevent a data controller from processing their data for particular purposes, such as by withdrawing consent or using their rights to object to or restrict processing. Individuals are also entitled to a range of metadata, including retention time, data source, purposes of processing, and meaningful logic about certain types of automated decision-making. These 'building blocks' might be used individually or in combination to facilitate enhanced transparency and control.

Many of these rights have existed in some form and in some jurisdictions since the 1970s [10] and became more widespread (both within Europe and outside) since the adoption

[2] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.

of laws based on the EU 1995 Data Protection Directive (DPD).[3] Despite their presence on the statute books of more and more countries, the rights, and associated obligations, have been subject to very little judicial scrutiny and there is considerable uncertainty about how the rules should be applied in specific circumstances, particularly concerning fast-moving technologies.

Furthermore, while theoretically strong, they tend to be burdensome for data subjects to use in practice, and are often ineffective as a result [7]. Many of these networked challenges are difficult to grapple with from an individual point of view, as even understanding the kinds of data out there in the world, the kinds of business relationships that exist, and the kinds of rights that are available, is challenging.

One major change in the GDPR from the DPD is the ability to submit requests electronically, and receive responses, including copies of the data, in commonly-used and/or machine-readable formats. This is a significant change from the *status quo*, where companies, even those with almost no offline presence for consumers, often attempted to force individuals to write a physical letter to their European headquarters, including a fee for the request (up to £10 in the UK), often requested through a cumbersome payment method such as a postal order rather than an online payment. These approaches, which may be employed to deter the exercise of individual rights, have now been more-or-less scrapped, meaning that requests can be made instantly and without cost.

In this working group, we discussed the possibilities and utility of electronic use of these rights, potentially through semi-automated means *de facto* delegated to a third-party. Delegating rights in this way might come with risks, such as to privacy, but also benefits, such as avoiding individualising the challenge of accountability. Furthermore, deploying emerging technologies may help balance these trade-offs. Some potential impacts and challenges are outlined below.

3.7.2.1 Information rights to understand provenance

Various information provisions in the GDPR allow a requester to learn where data came from (backward provenance), and to whom the data have been, or may be, transferred (forward provenance).

Article 13 specifies information that should be provided to the data subject at the point of collection. Article 13(1)(e) states that, when data are collected from a data subject, the data subject should be provided with forward provenance information: ‘the recipients or categories of recipients of the personal data, if any’.[4] Ideally, the information provided here would be sufficient to allow a data subject to query the recipients of these data, and further expand their enquiries. Unfortunately, the proposed text was contentious, and ultimately retained language from the DPD that the data subject might also be provided with the *categories* of recipients rather than specific named recipients. European regulators, in their guidance on the topic, have however emphasised that the default should be *named* recipients, and if a data controller wishes to use categories of recipients, they should justify why they believe this to be fair, and provide sufficiently detailed descriptions including, for example, a breakdown by geographical location and sub-sector [2].

[3] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ 1995 L 281/31.

[4] An individual can also obtain this information on request using Article 15(1)(c).

Learning where data came from (backward provenance) however has considerably less ambiguity compared to learning where it was going. Where data are not directly collected from a data subject, Article 14 applies, requiring similar information to be provided to the data subject as under Article 13. Where the individual is difficult to communicate with without disproportionate effort (for example, if re-identifiable but not identified data are transferred), such information might be published more generally, such as on a website. Article 14(2)(f) provides a backward provenance provision that states that 'the controller shall provide the data subject with the following information necessary to ensure fair and transparent processing in respect of the data subject [...] from which source the personal data originate, and if applicable, whether it came from publicly accessible sources'.[5] The *source* of the data is not qualified to give a data controller the option to substitute more general information, such as categories. As a result, the GDPR appears to be stronger in its provision for backward provenance than it is for forward provenance.

For the purposes of accountability, backward provenance has unfortunate downsides compared to forward provenance, as the data subject is more likely to be able to identify the data controllers that have collected data from them directly compared to those that have been passed data, potentially through long chains of interactions. Nevertheless, the GDPR provides for some transparency regarding both types of data collection, and automated requests can therefore be attempted automatically in relation to data provided both directly and indirectly.

If regulators and courts take this direction, Article 14 appears to oblige many firms to publish detailed information on where data came from. In that case, systems designed to scrape websites and privacy policies might be able to help trace back provenance trails and establish further oversight around the movement of personal data. Failing this, individual requests will likely remain necessary. The next section considers some of the socio-technical challenges associated with making such requests in automated and semi-automated ways.

3.7.3 Challenges

3.7.3.1 Identification and verification

A key challenge in all of this is being able to connect data accurately with an individual. In some databases, an individual might be easily recognised via an email address, phone number, or other obviously recognisable identifier. In other cases, an individual's identity may be inferred from information less obvious to the individual, such as an ID associated with a phone or other tracking device.

When logging into an online service, a user may be able to choose from among several IDs, such as a site-specific ID, a government ID, a Google ID, or a Facebook ID. Are these all equivalent, or is the user exposed to different risks depending on which identity is chosen? Is the identity service, like Google, used solely to authenticate access, or do other data leak between the identity service and the service being logged into? Can the average user be expected to know or understand these risks? It is also common to 'single users out' using information such as identifiers, cookies and web fingerprints which are difficult or even impossible for users to find or send, but comparatively easy for data controllers to use [5, 15].

To complicate matters further, devices and logins are often shared. A child might borrow a parent's phone. A couple might share an online subscription. The presence of a visiting guest might result in changes to electricity usage or TV watching patterns. In many cases

[5] Article 15 provides a similar provision upon request.

like these, it would be easy to be mistaken as to whom the data relate. As a result, there is often considerable uncertainty regarding the identity of the relevant data subject, yet this uncertainty is generally not acknowledged.

3.7.3.2 Dialogue with data controllers

Access requests are rarely smooth and simple to make. Recent research has highlighted that users face a difficult task in obtaining rights they are entitled to. Firms frequently point users to an automated data download tool when they request access to their data, even though such a tool is often non-exhaustive in terms of the categories of data it provides [3]. Similarly, requests can be confusing for those receiving them, or identification (see above) can be challenging. Smaller data controllers may simply lack the expertise or resources required to respond appropriately or fully to requests.

Such difficulties are similar to those experienced by those making freedom of information (FOI) requests to public bodies under the variety of FOI laws in place around the world. A range of tacit and codified knowledge has emerged around how to make successful FOI requests [6], and civic technology platforms such as *WhatDoTheyKnow.com* and *AskTheEU.org* have been set-up to help guide individuals through their ping-pong email exchanges with officials until they get the information they are seeking. The most successful and revealing FOI requests are frequently those in which the requester already knows a significant amount regarding the information that exists, such as the title, date or even the reference for a document [6]. This knowledge barrier has led FOI to become a tool-of-the-trade for journalists, but a right much less frequently utilised by laypersons.

Data protection rights have historically had a similar fate. While limited information is available on the use of data protection rights across sectors, given the lack of any reporting obligations for the majority of the organisations that are subject to the law, the evidence that exists points to a highly professionalised use of rights. The Law Society of England and Wales, for example, has highlighted the extensive use of subject access rights in immigration proceedings against organisations such as the UK Home Office [11]. For successful access requests in sectors with very different forms and categories of data, it seems likely that specific help to target and obtain desired outcomes will be required.

Many of these back-and-forth interactions will be undertaken via email, and there is a need for technologies to support individuals in asking the right questions, and having access to the relevant legal arguments that allow them to clarify the obligations of data controllers (who may not be aware of the extent of their obligations, such as the breadth of what is considered ‘personal data’ [12]) and to be aware when the arguments being made by data controllers may not have substantive legal backing, and could be grounds to complain to a data protection authority. Such technologies in other fields have been popular in recent years, such as the well-publicised chatbot *DoNotPay*, which helps individuals to overturn parking tickets [8]. Response prediction technologies have also entered common public use, such as Google’s *Smart Reply* feature, which predicts an appropriate short phrase on the basis of the content of a received message [9]. Yet given that data protection regulations are more complicated than parking violations (which can be managed via a relatively simple rule-based system), and that the aim is not to predict a realistic response like *Smart Reply* but to achieve a particular outcome, in many ways this issue is more challenging, and is reminiscent of work in legal expert systems.

Significant interdisciplinary research is required in this field. Such a platform to enable data subjects to utilise their rights must be legally sound and up-to-date with the latest case law, particularly as more cases are expected in the Court of Justice of the European

Union now that rights will be considerably easier to exercise. It must also be usable for data subjects; both easy to understand and providing practical assistance to them to achieve their desired outcomes. To enable this, it likely that a considerable amount of advanced natural language processing will be required.

The difficulty of developing a usable system to solve this complex problem is compounded by privacy concerns. The requests that an individual is making, and what they care about, are likely to be highly revealing for some, especially where they concern sensitive data such as health status, or data used to prove identification. Centralising these data for the purposes of improving the training of, for example, a machine learning text classifier is risky. Failing to combine such data, however, is likely to make it difficult to improve the system in question, and may make it impossible to develop a system with appropriate usability. It may be useful to draw upon privacy-preserving, federated machine learning techniques to build models and run tests privately 'at the edge', however deploying these technologies is also far from straightforward, and replicating all the functionality possible in centralised data systems efficiently and effectively is still a subject of heavy research.

3.7.3.3 Data cleaning

The data formats returned by the access and portability rights are likely to pose challenges to the effective automation of data protection provisions. The right of access specifies that when a request is made by electronic means, data should be returned in a commonly used, electronic form (Article 15(3)). The right of portability goes further than this to specify that data be returned in a "structured, commonly used and machine-readable format" (Article 20(1)), although the categories of data that can be requested through a portability request are limited in the text to those which a data subject has "provided to a controller" and based on consent or contract, unlike the right of access which concerns a wider "copy of the personal data undergoing processing" regardless of its lawful basis (Article 15(3)).[6]

One potential outcome is that many access requests will be fulfilled with a format such as PDF, which is notoriously difficult to extract information from reliably and without human intervention. Even seasoned data analysts find table extraction from formats such as PDF difficult, and a considerable research literature exists around how to do this reliably and repeatedly given the wide variety of ways in which data are presented [13]. Furthermore, structures of data which are readable easily by humans, such as 'wide' tables can be difficult to do structured analysis on with data manipulation grammar [16]. Transforming data between these formats can be a challenging task even for seasoned analysts, let alone for laypeople. As above, this data is likely to be highly sensitive, and support tools for transformation will be likely required to run locally with challenges in applying learning technologies to automatically process different controllers' datasets due to the difficulties in amassing sanitised examples.

3.7.3.4 Children

Children present a special class of user community. What data should be collected about children? How do children exercise their subject access rights, or can parents do this on their

[6] There is some controversy over this, as the pan-European group of regulators, the Article 29 Working Party has stated in its guidelines on portability that Article 20 refers to data that have been both "actively and knowingly provided by the data subject" and "[o]bserved data provided by the data subject by virtue of the use of the service or the device". Whether the Court will agree with the regulators on this apparent broadening of the original text remains to be seen. See [1].

behalf? Some "competent" children may be able to make requests on their own behalf, but how is competency determined?

Consent in GDPR is country-specific. In many countries, 16 is considered the minimum age for consent but there is considerable deviation from this norm, with the age ranging as low as 13.[7] In general, GDPR is not clear about children exercising rights, so one needs to look at non-binding recitals (e.g. recital 65) for guidance and not just the GDPR's core text (the articles).

The text and guidance do not anticipate children exercising rights directly, and yet these are among the most vulnerable groups for whom we might want to ensure strong accountability and oversight. The ICO has emphasised that data controllers should "allow competent children to exercise their own data protection rights."[8]. However, it is important to consider how competence is measured, both in a legal sense as well as the potential risks that might be present from giving children access to all their personal data which they might then be pressured to give away to companies or other parties.

3.7.3.5 Additional issues

Additional issues that were discussed include

- how to ensure the data requested under these rights does not become a security risk in and of itself when held locally by the users;
- how data controllers might be empowered to ensure that automated data rights do not cripple the digital economy (e.g. what sort of APIs should be promoted?);
- what kind of analysis of requested data might be undertaken with which methods, how revealing it might be and how useful it might be to a regulator;
- whether authentication methods might be possible using zero-knowledge proofs or similar approaches;
- the roles of third party intermediaries (such as technology giants) in providing 'scalable access request management', and how this might affect privacy and competition, and whether it might be possible to undertake using secure cloud and enclave approaches;
- standardisation efforts for data portability and access.

3.7.4 Conclusions and open questions

Many of the open questions in this field will only be dealt with by trying to build systems that attempt to automate data rights. To what extent those systems will find support in the law, such as in provisions around data protection by design, remains unclear. Furthermore, some controllers might see such systems as a threat to trade secrets and confidentiality, and may take protective measures to prevent the aggregation of release. In other sectors, standardisation might be welcomed for the purposes of economic efficiency and reduced compliance burden.

In the short term, we identified needs for

- An ongoing meta study of all the GDPR access request research that will be written after everyone starts collecting data on 25 May;

[7] Mapping the GDPR age of consent across the EU: April 2018 update at https://www.betterinternetforkids.eu/web/portal/practice/awareness/detail?articleId=3017751

[8] https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/applications/children/

- Estimations of the rates of chaos from email-based systems at scale to provide evidence and support for standardisation and automation from the data controllers' perspectives;
- Connection to research in personal data containers such as DataBox, openPDS, the Hub-of-All-Things, among others;
- Usability research into how individuals grapple with their own rights, what their personal goals are and what support tools they need and find useful;
- Local and sectoral differences in subject access request fulfilment and enforcement;
- The role of identity providers and verification, working with regulators to ascertain appropriate security-data rights trade-offs and to implement them technically in different contexts;
- The structure of an open source personal data management system founded on automated rights: what core technologies (e.g. privacy-enhancing technologies) are best suited, and which open research questions in those domains are a pre-requisite to some of the aims (such as analytics at scale for systemic accountability) we have here.

References

1 Article 29 Data Protection Working Party. *Guidelines on the right to "data portability" (wp242rev.01).* 2017.

2 Article 29 Data Protection Working Party. *Guidelines on Transparency under Regulation 2016/679 (wp260rev.01).* 2018.

3 Jef Ausloos and Pierre Dewitte. Shattering one-way mirrors – data subject access rights in practice. *International Data Privacy Law*, 8(1):4–28, February 2018.

4 Reuben Binns, Ulrik Lyngs, Max Van Kleek, Jun Zhao, Timothy Libert, and Nigel Shadbolt. Third party tracking in the mobile ecosystem. In *Proceedings of the 10th ACM Conference on Web Science*, pages 23–31. ACM, 2018.

5 Frederik J Zuiderveen Borgesius. Singling out people without knowing their names–behavioural targeting, pseudonymous data, and the new data protection regulation. *Computer Law & Security Review*, 32(2):256–271, 2016.

6 Matthew Burgess. *Freedom of Information: A Practical Guide for UK Journalists.* Routledge, June 2015.

7 Lilian Edwards and Michael Veale. Slave to the Algorithm? Why a 'Right to an Explanation' is Probably Not The Remedy You Are Looking For. *Duke Law Technol. Rev.*, 16(1):18–84, 2017.

8 Samuel Gibbs. Chatbot lawyer overturns 160,000 parking tickets in London and New York. *The Guardian*, June 2016.

9 Anjuli Kannan, Karol Kurach, Sujith Ravi, Tobias Kaufmann, Andrew Tomkins, Balint Miklos, Greg Corrado, Laszlo Lukacs, Marina Ganea, Peter Young, and Vivek Ramavajjala. Smart reply: Automated response suggestion for email. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 955–964, New York, NY, USA, 2016. ACM.

10 Christopher J Millard. *Legal protection of computer programs and data.* Sweet & Maxwell, Toronto and London, 1985.

11 The Law Society of England and Wales. *Written evidence from the Law Society of England and Wales to the House of Commons Public Bill Committee considering the Data Protection Bill.* 2017.

12 Nadezhda Purtova. The law of everything. broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, pages 1–42, April 2018.

13 Roya Rastan, Hye-Young Paik, and John Shepherd. TEXUS: A task-based approach for table extraction and understanding. In *Proceedings of the 2015 ACM Symposium on Document Engineering*, pages 25–34. ACM, September 2015.

14 Narseo Vallina-Rodriguez, Srikanth Sundaresan, Abbas Razaghpanah, Rishab Nithyanand, Mark Allman, Christian Kreibich, and Phillipa Gill. Tracking the trackers: Towards understanding the mobile advertising and tracking ecosystem. *arXiv preprint arXiv:1609.07190*, 2016.
15 Michael Veale, Reuben Binns, and Jef Ausloos. When data protection by design and data subject rights clash. *International Data Privacy Law*, 8(2):105–123, 2018.
16 Hadley Wickham. Tidy data. *J. Stat. Softw.*, 59(1):1–23, 2014.

4 Overview of Talks

4.1 Tutorial–Cloudy with a hint of accountability

Jon Crowcroft (University of Cambridge, GB)

Joint work of Haddadi Mortier et al.
Main reference Richard Mortier, Jianxin R. Zhao, Jon Crowcroft, Liang Wang, Qi Li, Hamed Haddadi, Yousef Amar, Andy Crabtree, James A. Colley, Tom Lodge, Tosh Brown, Derek McAuley, Chris Greenhalgh: "Personal Data Management with the Databox: What's Inside the Box?", in Proc. of the 2016 ACM Workshop on Cloud-Assisted Networking, CAN@CoNEXT 2016, Irvine, California, USA, December 12, 2016, pp. 49–54, ACM, 2016.
URL http://dx.doi.org/10.1145/3010079.3010082

This was an introductory talk about the use of different approaches to improve the confidentiality and integrity of cloud based systems. On the one hand, in large data centers, we can use Trusted Execution Platforms running on newer hardware (both SGX on Intel CPUs and Trustzone on Arm processors) to provide much stronger assurances about the privacy, and attest to integrity of code and data processing in the cloud. On the other hand, we can use edge-cloud (or personal cloud) to run systems in the same way, but keeping data and processing in its home (smart home, car, building city, etc.) improving availability, latency, energy efficiency, and reducing the attackable footprint ("surface") of the system. Linking this with accountability is a work for the future. Of course, systems can still shoot themselves in their feet.

4.2 Tutorial–ML : transparency, control, user rights and the GDPR

Lilian Edwards (The University of Strathclyde – Glasgow, GB)

Main reference Lilian Edwards and Michael Veale: "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For" 16 Duke Law & Technology Review 18 (2017)
URL https://dx.doi.org/10.2139/ssrn.2972855

The tutorial outlined the basic principles and structure of the General Data Protection Regulation (GDPR). The conditions for lawful processing, including, but not restricted to, consent, were outlined as were the definitions and import of the terms personal data, processing and data controller/data processor. The rest of the tutorial concentrated on the DP problems that arise out of "Big Data", profiling and machine learning. Particular attention was paid to the so-called "right to an explanation" which might be derived from art 22 or art 15(h); and issues arising from its implementation in EU member states.

4.3 Tutorial–A Brief Introduction to Provenance in Workflows and Databases

Bertram Ludäscher (University of Illinois at Urbana-Champaign, US)

Computational notions of data provenance have been studied in different contexts such as databases, programming languages, and scientific workflows. While there are considerable differences in the assumptions, perspectives, and problems studied by the various communities, a common underlying theme is that of transparency and comprehensibility of the computational processes that yield a given data output. When trying to make complex systems "accountable" for data-driven, algorithmic actions and decisions, it is necessary to capture relevant provenance information, e.g., by "looking inside" of black-box computations and capturing not only retrospective provenance but also prospective provenance, i.e., the dataflow dependencies of the computations and workflows that make up the computational system. Hybrid forms of provenance, combining workflow specifications and retrospective provenance provide new opportunities to gain insights into the intended and actual workflow executions, and can be used to account for and explain system behavior. New research challenges arise from the inherent trade-off between transparency and provenance (necessary elements of accountability) on one hand, and requirements for privacy and data protection on the other.

References

1 Ludäscher, Bertram. A Brief Tour Through Provenance in Scientific Workflows and Databases. In *Building Trust in Information*, edited by Victoria L. Lemieux, 103–26. Springer Proceedings in Business and Economics, 2016.

2 Dey, Saumen C., Daniel Zinn, and Bertram Ludšcher. Propub: Towards a declarative approach for publishing customized, policy-aware provenance. In *International Conference on Scientific and Statistical Database Management (SSDBM)*, pp. 225–243. Springer, Berlin, Heidelberg, 2011.

4.4 Tutorial–Ethical/social: Rights, Ethics and Accountability

Ben Wagner (Wirtschaftsuniversität Wien, AT)

Main reference Ben Wagner, "Algorithms and Human Rights: Study on the human rights dimensions of automated data processing techniques and possible regulatory implications" (No. DGI(2017)12). Strasbourg, France: Council of Europe.
URL https://rm.coe.int/algorithms-and-human-rights-study-on-the-human-rights-dimension-of-aut/1680796d10

The debate about ethical and social dimensions of automated systems is frequently limited to bias. The following talk provides an overview of the debate around automation/algorithms/AI in information systems and the confusion around concepts of ethical and regulatory solutions within it. It suggests that there is a need to answer key questions about algorithmic accountability are answered: accountable to whom, where and what for. It also suggests some key shifts in the debate to ensure meaningful accountability rather than just a fig leaf. The talk is based on the two publications listed below.

References

1 Ben Wagner. "Algorithms and Human Rights: Study on the human rights dimensions of automated data processing techniques and possible regulatory implications (No. DGI(2017)12)." Council of Europe, 2018 https://rm.coe.int/algorithms-and-human-rights-study-on-the-human-rights-dimension-of-aut/1680796d10.

2 Ben Wagner. "Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?" In *M. Hildebrandt (Ed.), Being Profiling. Cogitas ergo sum.* Amsterdam University Press, 2018.

4.5 Raising Users' Awareness for their Exposure to Cloud Services

Martin Henze (RWTH Aachen, DE)

Main reference Martin Henze, Jan Pennekamp, David Hellmanns, Erik Mühmer, Jan Henrik Ziegeldorf, Arthur Drichel, Klaus Wehrle: "CloudAnalyzer: Uncovering the Cloud Usage of Mobile Apps," in Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous), ACM, 2017.
URL http://dx.doi.org/10.1145/3144457.3144471

Users are often unaware of their exposure to cloud services, e.g., when sending and receiving emails or when interacting with mobile apps on their smartphones. However, only if users are aware of (the extent of) their exposure to cloud services, they can make informed decisions and exercise their right to privacy. As a foundation to put users back into control over their privacy, we hence consider it necessary to uncover their exposure to cloud services and raise their awareness of resulting privacy risks. In this talk, we present approaches to provide users with transparency over their individual exposure to cloud services along two deployment domains for cloud services even less technically proficient users interact with on a daily basis: email and mobile apps on smartphones. Furthermore, we discuss how to apply the concept of comparison-based privacy to enable users to put their cloud usage into context through comparison with their peers in a privacy-preserving manner.

References

1 Martin Henze, Ritsuma Inaba, Ina Berenice Fink, and Jan Henrik Ziegeldorf. Privacy-preserving Comparison of Cloud Exposure Induced by Mobile Apps. In *Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous)*. ACM, 2017.

2 Martin Henze, Jan Pennekamp, David Hellmanns, Erik Mühmer, Jan Henrik Ziegeldorf, Arthur Drichel, and Klaus Wehrle. CloudAnalyzer: Uncovering the Cloud Usage of Mobile Apps. In *Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous)*. ACM, 2017.

3 Martin Henze, Mary Peyton Sanford, and Oliver Hohlfeld. Veiled in Clouds? Assessing the Prevalence of Cloud Computing in the Email Landscape. In *Proceedings of the 2017 Network Traffic Measurement and Analysis Conference (TMA)*. IEEE/IFIP, 2017.

4.6 Purpose-driven provenance solutions

Melanie Herschel (Universität Stuttgart, DE)

Provenance in general is meta-data about the production process of an end product, including digital data. Depending on the intended purpose, the type of process, the underlying data, and the processing environment, different provenance solutions to capture, store, and use provenance are conceivable. This talk argues on the necessity to design, adapt and optimize provenance solutions to specific environments. We briefly present two examples of such purpose-driven provenance solutions: provenance for debugging data processing to account for processing errors or ensure more robust processing and provenance used to document and analyze visual computing program behavior.

References

1 Melanie Herschel, Ralf Diestelkämper, Houssem Ben Lahmar: A survey on provenance: What for? What form? What from? *VLDB Journal*, 26(6): 881–906 (2017)

4.7 Accountable systems. Some practical experiences from a governmental perspective.

Heleen Louise Janssen (Ministry of the Interior and Kingdom Relations, NL)

How are constitutional law and digital technologies being reconciled from an ‘on the ground perspective’? Focus in this lightning talk is on fundamental rights. Fundamental rights are supposed to be solid, at least for a generation. Because of their broad application and large time-span they are inherently vague. Fundamental rights apply in principle solely to public entities, but may still have impact on private relationships (e.g. between companies and consumers or users) via ordinary legislation or in the explanation of tort laws in court cases.

In my advisory practice from the Constitutional Affairs and Legislation Department concerning the use of technological tools with the aim to execute policy, we strongly recommend that fundamental rights issues are already being solved in the design stage. Artefacts have politics–technological applications are value laden.

As regards the reconciliation of constitutional legal reality and technological reality, the modernisation of the right to communication secrecy was presented. The constitution (dating from 1983) only protects content that is transported or transmitted by physical means of mail, telephone and fax (a closed list of means). The modernisation entailed a broadening to all possible means transporting or transmitting content, covering also the electronic means of communication to make it future-proof. Various technological and legal issues had to be reconciled. Is content that is transmitted via WhatsApp or Facebook being protected, are they ‘means’ like a telephone or a letter? Why is an identifiable receiver important in a world where communication can be multiplied a billion times? When does meta-data (from a technological perspective not content) come close to or even become (legal) content, and what are the consequences for its legal protection? Is content transmitted in IoT protected by communication secrecy?

In the second example about the modernisation of the Intelligence and Security Services Act (ISSA), comparable issues from a fundamental rights perspective had to be solved. Issues dealt with in the context of privacy and communication secrecy concerned the question whether ‘scraping’ the internet as an open source would invade privacy. Another question that had to be solved was how the standing big data processing methods could be reconciled with data protection requirements such as purpose limitation and data minimisation. The law was presented in an internet consultation (1100 reactions) and a Privacy Impact Assessment (done by independent researchers) that was presented with the Bill to Parliament.

Research undertaken by the University of Utrecht (NL) at the request of the Ministry of the Interior has demonstrated that not only privacy, but also the right to equal treatment and procedural rights (access to court, right an effective remedy) are under severe pressure due to the application of big data analytics, IoT and AI.

4.8 Establishing Requirements for Accountable Systems: The Case of Elections

Joshua A. Kroll (University of California – Berkeley, US)

In setting requirements for designing a system, designers must consider how the system’s mechanisms support its high-level goals. These goals can include the reflection of important human values, such as accountability or consistency with social, political, and legal norms. It is also important to hold the process that sets these requirements accountable so as to minimize the gap between the requirements as desired by stakeholders and the requirements as they are articulated or realized in the final fielded system.

We consider the case of designing an election system, which has received much attention in the security literature [2, 1]. Elections provide an excellent test case for thinking about design that supports accountability and other values. In addition to the functional requirement of registering and tallying voter intent, election systems must attest convincingly to the integrity of results even in difficult, adversarial political and security environments, and often need to do so while maintaining the privacy of ballot information.

To accomplish this, election systems must track eligible voters from registration through casting and also maintain a strong and verifiable chain-of-custody on ballot materials from ballot design through audits that may come after the certification of results. Additionally, election systems need to be sensitive to usability for all voters, to avoid inadvertently disenfranchising certain sub-populations. Election systems must also be highly available and resilient to many types of failure so as to meet the requirement that they capture voter intent reliably within a short, fixed window of time. Post-election audits can be used to increase confidence in a result or to challenge the outcome. At every stage, the system must generate sufficient detailed evidence of correct operation so that subsequent challenges do not undermine the legitimacy of the announced outcome.

References

1 Joseph Lorenzo Hall. *Policy mechanisms for increasing transparency in electronic voting.* PhD thesis, 2008.

2 Joseph Lorenzo Hall. Election auditing bibliography, 2010. https://josephhall.org/papers/auditing_biblio.pdf.

4.9 Confidential Analytics – Use Cases and Building Blocks

Maximilian Ott (CSIRO – Alexandria, AU)

URL https://n1analytics.com

Data collaboration between organisations using sensitive data is not only increasing in volume but also in the problems it creates. While there are extremely interesting ethical questions raised about the intent of many of those collaborations, most of today's problems are caused by the "process" leading up to the desired outcome.

Specifically, today's analytics tools require that all the input data is brought together in a single place, requiring at least one party to "disclose" potentially sensitive data.

In the first part of this talk, we describe a few use cases for data collaboration. We then briefly sketch an alternative analytics approach where all the sensitive data can remain with the data owner. This is followed by a short description of the required building blocks: arithmetic with encrypted numbers, federated algorithms and protocols, as well as private record linkage.

4.10 Accountable systems, messy environments

Michael Veale (University College London, GB)

Joint work of Michael Veale, Max Van Kleek, Reuben Binns, Lilian Edwards
Main reference Michael Veale, Max Van Kleek, Reuben Binns: "Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making", in Proc. of the 2018 CHI Conference on Human Factors in Computing Systems, CHI 2018, Montreal, QC, Canada, April 21-26, 2018, p. 440, ACM, 2018.
URL http://dx.doi.org/10.1145/3173574.3174014

Responding to concerns around 'algorithmic harms', fields such as law and computer science have doubled down on efforts to create facilitating technologies and environments for accountable systems. Canonical problems include biased, opaque automated decision-making systems, and canonical solutions include discrimination-aware machine learning, user-facing explanation facilities, or data protection information rights. In this talk, I present two areas where these canonical visions of these challenges clash with practice. Firstly, some results from a study interviewing 27 public sector machine learning practitioners is presented, where their attempts to cope with accountability issues are shown to be more textured in practice than commonly suspected. Secondly, I consider the information rights in the EU General Data Protection Regulation, and how they might be able to support provenance and transparency efforts, and may be promising at scale or where used collectively, but are considerably messier in practice than the text would lead a reader to believe.

4.11 Trust, Responsibility, and Explanation

Michael Winikoff (University of Otago, NZ)

Joint work of Virginia Dignum, Frank Dignum, Michael Winikoff

My talk considered the overarching issue of trusting autonomous systems, and the factors that lead to appropriate levels of trust in autonomous systems. I particularly focussed on the role of explanation, and described an explanation mechanism and its evaluation.

Participants

- Virgilio Almeida
Federal University of Minas Gerais-Belo Horizonte, BR
- Jean Bacon
University of Cambridge, GB
- Jennifer Cobbe
University of Cambridge, GB
- Jon Crowcroft
University of Cambridge, GB
- Lilian Edwards
The University of Strathclyde – Glasgow, GB
- David Eyers
University of Otago, NZ
- Krishna P. Gummadi
MPI-SWS – Saarbrücken, DE
- Tristan Henderson
University of St Andrews, GB
- Martin Henze
RWTH Aachen, DE
- Melanie Herschel
Universität Stuttgart, DE
- Heleen Louise Janssen
Ministry of the Interior and Kindom Relations, NL
- Joshua A. Kroll
University of California – Berkeley, US
- Bertram Ludäscher
University of Illinois at Urbana-Champaign, US
- Derek McAuley
University of Nottingham, GB
- Christopher Millard
Queen Mary University of London, GB
- Ken Moody
University of Cambridge, GB
- Maximilian Ott
CSIRO – Alexandria, AU
- Frank Pallas
TU Berlin, DE
- Thomas Pasquier
University of Cambridge, GB
- Silvia Puglisi
The Tor Project – Seattle, US
- Margo Seltzer
Harvard University – Cambridge, US
- Jatinder Singh
University of Cambridge, GB
- Barbara Staudt Lerner
Mount Holyoke College – South Hadley, US
- Michael Veale
University College London, GB
- Ben Wagner
Wirtschaftsuniversität Wien, AT
- Michael Winikoff
University of Otago, NZ
- Martina Zitterbart
KIT – Karlsruher Institut für Technologie, DE