Report of the Internet Privacy Engineering Network (IPEN) Workshop:
'Engineering Privacy into Internet Services and Applications'[1]

Berlin / 26 September 2014

The first Internet Privacy Engineering Network (IPEN) Workshop was held at the Berlin State Parliament (Abgeordnetenhaus von Berlin) on 26 September 2014. The workshop, entitled 'Engineering Privacy into Internet Services and Applications', was co-hosted by the Berlin DPA, the European Academy for Information Freedom & Data Protection (EAID), the OWASP Privacy Risks Project, Oxford Internet Institute, University College London, and the DPAs of Berlin, France, Ireland, the Netherlands, the United Kingdom, Schleswig-Holstein (DE) and the EDPS.

More than 55 participants took part in round table discussions on the development of technical solutions to the most pressing privacy engineering needs. Keynote speakers included Peter Hustinx (European Data Protection Supervisor), Dr. Alexander Dix (Berlin Commissioner for Data Protection and Freedom of Information) and Peter Schaar (Chairman of the European Academy for Information Freedom & Data Protection - EAID).

This is a report of the discussions and the themes which merit further work. Anyone interested in further information or to contribute views and expertise to this project is welcome to contact us at ipen@edps.europa.eu

Workshop sessions can also be viewed on-line here:
https://www.youtube.com/watch?v=8alKHBh7tN4&list=PLgrQeHXIMX5O5D0KPIIyuTTHzjSZVraIJ.

Keynotes

The President of the Berlin Parliament, Mr **Ralf Wieland**, opened the workshop and welcomed the participants, introducing them to the historical building of the former Prussian parliament. The building was the place of key events in recent German history and where the importance of democracy and civil liberties could be witnessed.

Peter Hustinx noted that the workshop, while not the first conference on privacy and technology, was the first European meeting which took this very practical approach in the engineering of privacy on the internet.

Although legislators and privacy regulators have been investing considerably in defining and interpreting rules for the protection of personal data and privacy, new Internet tools and applications are still often implemented and deployed without proper data protection safeguards. There is significant risk that this gap will increase. Augmenting the risk to individuals' privacy and fostering embedded surveillance instead of embedded privacy. This process could invalidate many of the efforts of data protection authorities.

[1] Disclaimer: This is a report of the discussions at the workshop. It represents statements made by the participants and does not constitute a position taken by the EDPS.

Last year's revelations about massive communications surveillance have considerably raised awareness about the huge risks for confidentiality and security of personal information. This experience has been an alarm for society, and led to the decision by internet engineers to treat mass surveillance as a technical challenge.

The ultimate purpose of the IPEN initiative is to narrow the gap between available technical tools and best practices on one hand and privacy needs on the other hand. It should encourage developers to create privacy friendly tools. The first step in this endeavour is to improve communication between privacy experts and the developer communities (business, free software engineers and academia).

Dr. Alexander Dix stated that neither law, nor technology alone offer solutions to privacy issues, in particular, it is unlikely that there will ever be a privacy button which one can press and receive privacy. "We are at a crossroads and need usable and effective tools for informational self-defence." Dix stated that we need cryptographic solutions, but that cryptography and encryption are not enough. Tool boxes for service providers, manufactures and users are needed.

Peter Schaar noted the need for networking among the multiple actors in the data protection field: Data Protection Authorities, civil society, industry, and academia. IPEN was created for continuous cooperation in privacy engineering. Data security and data protection are strongly related, and much of the data surveillance procedures are not visible to internet users. While data protection is traditionally seen as a legal question, legal frameworks are not enough. Rather, robust structures, services, technical platforms can provide fortification in cases where legal systems alone cannot guarantee data protection.

Session 1: Exploring the existing initiatives & tools and identifying the technical gaps.

The experts from the **OWASP Privacy Risks Project**[2] established a rating of the top ten privacy issues considering both technical and organisations risks. With roughly 150 experts involved, 'web application vulnerabilities' is the top issue on the list. It is an expert survey rating, i.e., it is not based on statistical data, but rather on the perceptions of the participating experts.

Approaches and standards developed by the **Internet Engineering Task Forces (IETF), by Internet Architecture Board (IAB) and by the WorldWideWeb Consortium (W3C)** Privacy Interest Group were briefly described. While the IETF has established a review process for new standards, this initially focused on security as a design consideration. It is not presently possible to design and standardise new protocols without confidentiality, authentication, integrity, and similar protections, unless there are strong arguments to support this choice. However in this approach, privacy elements are weak and the terminology is confusing. That is why the Internet Architecture Board (IAB) created guidance which extends privacy considerations to a broader definition than security alone, for example, its RFC 6973, but is not a mandate. Furthermore, the W3C Privacy Interest Group (with members of ISOC and Apple) drafted a privacy consideration document. All these activities prove that important efforts have been made in the privacy field. Hopefully, these efforts will be

[2] https://www.owasp.org/index.php/OWASP_Top_10_Privacy_Risks_Project#tab=Top_10_Privacy_Risks

accompanied by an increased awareness within the standards community in terms of privacy and technology. Furthermore, what is needed to follow is attention to the phases of implementation and deployment and to provide guidance there.

The Tor software is an example of a software tool developed with the objective of providing privacy to Internet users. TorServers.net operates Tor servers worldwide to provide means for secure and anonymous communication.

Tor's original design was made more than 12 years ago. It relies on directory authorities which know about every Tor relay in the network. A Tor client downloads a list of all relays and chooses three servers. It builds an encrypted channel and bounces traffic back and forth. In the early days of Tor the software had to be set up as a proxy in the browser. However some browsers leaked data, and it was necessary to develop a secure variant. Nowadays the Tor Browser is a modified version of Mozilla's Firefox. It comes with patches to enhance privacy and security and uses Tor to transmit data.

Bridges are non-visible entry points into the Tor network. There is now an application for Android mobile phone Tor use called OrBot, Onioo measures censorship events and Tor Metrics shows current state of the Tor network.

Business architects are often left out of the privacy dialogue. As well, the use of a standardised terminology is lacking. From the legal point of view in the EU, data protection touches on human rights. From an engineering point of view, it is the nature or use of the data within a context that is important. Data itself is neutral; rather, it is how the data is used; and it is the involvement of the business architect that brings meaning to data. Privacy Impact Assessments identify risks or potential misuse of data; hence, analysing data flow is important. Article 29 Working Party provides useful opinions on privacy principles and practical use cases, such as: purpose limitation, anonymisation techniques, legitimate interests, application of necessity and proportionality.

Gaps in models is usually where errors occur. There is a difference between the gathering of personal data by corporations and the gathering of data by governments; however, NSA programmes show that state surveillance now largely relies on harvesting data from corporations. James Whitman, a Yale scholar, examines the differing EU and US conceptions of privacy in legal context. One gap in privacy engineering can be described as: 'pushing up the stack'. This describes the avoidance of privacy until later phases of development at higher levels, so that privacy protections are not built into the lowest layers of the design or implementation. This results in privacy concerns ultimately being enforced by technology or usage policies, rather than being built into systems.

Two experiments: Two experiments with two different aims were undertaken by the French Data Protection Authority, CNIL, along with the French National Institute for Research in Computer Science and Control (INRIA). Prior to giving pertinent recommendations and instructions, regulators must have a comprehensive understanding of the technical side. That is why the applications CookieViz and Mobilitics were created. (1) CookieViz is a visualization tool that measures the impact of cookies when you browse the Internet, or a type of Collusion or Lightbeam that works for all browsers. (2) Mobilitics provides in depth analyses of all personal information which is recorded,

stored, and disseminated by the smartphone. Its aim is also to foster further innovations and new sustainable services. There is the risk that certain companies hold so much data on individuals (e.g., Google's PlayStore). Also, many devices and operating systems link to users' locations, Unique Device ID (UDID) and device names (e.g., 16% of apps access device names).

Is it possible to entirely avoid identification? Sometimes it is not. Certainly there will be a much stronger push towards the services that allow a reasonable level of anonymity. But there has to be an incentive. Relying on regulation approaches or market based approaches to privacy is something we need to take. We need the regulation to force the adoption of the privacy friendly technologies.

Privacy risks related to the Internet of Things will have to be addressed.

Session 2: Use Cases - How can we identify and address privacy gaps?

Management of user rights: DPA's role with respect for PETs

It is always necessary to think how the data subject can exercise their rights, or the controllers - on behalf of data subjects. We need valid consent management, strict purpose limitations and strong security measures. Device fingerprinting is not inevitable, and it is not lawful to track by default. Why should we 'opt out'? Do end users really need a PhD in online advertising to be able to use their devices, exercise and protect their privacy? Technology needs to be at the service of data protection and not the reverse; therefore, we need privacy protocols. We have also to acknowledge that the technology left alone cannot solve everything; we need strict collaboration between the technology side and the policy/regulation side, pinpointed by standards.

In terms of describing the broad areas of concern that we come across currently where a technology to date protection gap may exist, access requests to personal data are among the top queries to a DPA. These may not be technological, but in some cases can be caused by a lack of definition or recognition by a Data Controller of what is and is not personal data. While unsolicited marketing is one of the limited issues where our courts have fined, our investigative and audit powers can produce positive data protection results sometimes over and above monetary fines. When problems come to light, we sometimes find that Data Controllers use data for purposes beyond the reasons described when initially collected, perhaps inadvertently or perhaps because it has not been considered a priority - especially for start-ups. Perhaps this might be as a result of a legal-technical divide, a communications gap, or a lack of a wider shared understanding between business, legal and technology. This gap between the binary technology world and the linguistic legal worlds can be evident for us in compliance terms also. The terms "appropriate", "reasonable" and "necessary" are difficult to translate and apply from regulation to software design, and for some legal experts, this can be an opportunity to earn their keep! In other areas, Big Data has the potential to be a large sector where personal data could be abused both by unbounded creative, innovative management and engineering people. It is an area that needs to be closely monitored so that it can be positively and properly harnessed and leveraged for good. Finally, it is not all about Data Controllers. Individuals play a part in managing their personal data and need to become more aware and consider their actions more closely, particularly when using 'freemium services'. However, that is not

to say that they are solely responsible for their own data protection especially when privacy invasion is not visible, making the user's job of discernment very difficult.

Usability of technology and the relationship with trust

For regular internet users, open source is too technical: (1st challenge) how to establish trust? (2nd challenge) how to address the profilers and data collectors? People from profiling businesses should be a part of this discussion as well. It is not only about the technology, but also about the incentives and business models. How do we address that? How to do proper 'privacy by design' for surveillance projects? Currently, privacy by design is simply a fig leaf in many surveillance projects. (3rd challenge): How to get privacy by design into practice? How to 'design privacy by design'? The real gap is how do we make software engineers design applications that respect privacy? How can we help them? Privacy by design will never be adopted unless engineers find it makes their lives easier. Moreover, web applications are generally built from existing frameworks, but the existing framework code is not supportive of data protection.

Regulation cannot be relied upon solely. Users are not in control of the internet or their interactions. Innovation at the edges of the internet is important and will continue to happen; therefore, privacy will always be an issue. Do not aim to solve the privacy problem, but aim to mitigate it. Another gap is the manner in which DPAs consider data v. metadata. Metadata can be equally informing. Much of the security technology has been too difficult to deploy at scale. Data minimisation is another gap, as it is often difficult to know how to achieve, as is usually not considered at early design phases. Privacy by design will never be designed if it does not make engineers lives easier. Produces a design methodology that is clear and improves the developers' work and processes. Lastly, there will always be bad actors: states, corporations, etc. More communications is needed between lawyers and engineers.

Privacy invasion is not necessary to conceive new technologies. Thus far, we have not had privacy by design, but rather surveillance by design. States assured that security was weak, e.g., telecom, Wi-Fi, and GSM protocols. Privacy is a public good, but the average user has no way to determine the levels or quality of privacy provided by products: i.e., a market failure. IPEN is 30 years late, but there are a few success stories: e.g., the option to turn on SSL or TLS encryption in emails (a large fraction of people chose to turn on SSL on Facebook). Users cannot be expected to know what is going on at the back end, and users are not to blame for this. The user experience is likely designed to circumvent their knowledge. The average application developers also cannot be expected to understand at the theoretical computer science level. Infrastructure, basic protocols and frameworks must support data protection. Most of the code for web apps is already written and not supportive of data subject rights. Infrastructure and application frameworks are the places to focus for better data protection.

The new Data Protection Regulation[3] is expected to be adopted in the course of 2015. The new Regulation will provide stronger rights for data subjects, stronger responsibilities for data

[3] On 25 January 2012, the European Commission adopted a package for reforming the European data protection framework. The package includes (i) a 'Communication' (COM(2012)9 final), (ii) a proposal for a 'General Data Protection Regulation' (COM(2012)11 final), and (iii) a proposal for a Directive on data protection in the area of criminal law enforcement (COM(2012)10 final). Two years after the publication of the Commission proposals, the Council of the

controllers, and stronger powers for enforcement by DPAs, but it will not address technology developers directly. It is implied that due to the new rules, there will be much stronger incentives, both positive and negative, for the creation of privacy friendly tools. However, is there a way to speed up this process for the demand of privacy compliant services?

Session 3: Approaches to engineering privacy

First conversation: Privacy considerations for Internet Protocols and Processing from a Societal level / Privacy Risk Assessments

Privacy considerations for Internet Protocols

Companies have a wide range of privacy principles to choose from, such as the FIPPS, or the Madrid Resolution. While these principles are sound, they are intentionally kept generic; and therefore, they do not provide sufficient guidance for engineering teams. An interpretation of these principles against the ever changing technology landscape is necessary. Unfortunately, the technology design process is quite complex and distributed, which also distributes responsibilities for introducing privacy features into various technical building blocks (such as APIs, protocols, algorithm implementations and libraries).

Before introducing changes to technology, however, it is important to raise the level of awareness among engineers and to illustrate how privacy influences the design decision process. For this purpose RFC 6973 was written, which introduces terminology, a list of privacy and security related threats, threat mitigation techniques, and guidance for protocol designers. In addition to the document, various tutorials have been held to reach a wider engineering audience. Seeking document review for upcoming standards (based on RFC 6973) turned out to be more difficult than expected since skilled and people willing to give feedback were hard to find.

With RFC 7258 “Pervasive Monitoring Is an Attack” a follow-up publication was created in response to the Snowden revelations, which calls for the increase of deploying of end-to-end security mechanisms.

These publications have received widespread attention among the technical community since it became apparent that increased privacy protection (for example due to the always-on encryption) impacts existing business models. This includes for example deep packet inspection for traffic management, SPAM filtering, and analytics. As the design of technology impacts business models and change the larger eco-system, the discussion about how to apply various privacy techniques is ongoing.

There is a need to raise levels of awareness among the technical community with privacy tutorials. Additionally, there are conflicts with existing business models (e.g., HTTP and use of proxies). Corporate frustrations can also result when 'geeks destroy business models'. There generally is no

European Union is to find a common position on the basis of which negotiations with the European Parliament and the Commission may begin.

privacy guidance provided at the implementation or deployment phases, which also needs improvement.

It was addressed that there are three (3) areas where engineers can best contribute in the field of privacy engineering: (1) in areas where enough building blocks already existing, (2) where impact in the network is significant, and (3) in areas that data protection authorities can verify and enforce personal data protection.

Processing from a Societal level / Privacy Risk Assessments

The task of developers and policy makers is to identify high impact issues. Therefore, it is important to analyse how people use technology in contrast to how developers envision the technology to be used. Also, it is important to analyse how the data and metadata is used. Metadata may, in certain situations, express more than the actual data itself. Legal terms such as 'legitimacy', 'data minimisation', 'purpose limitation', etc. have no real meaning to engineers. In contrast, the principle of 'transparency' can be addressed by engineers, for example, in the situation of device fingerprinting.

It is not enough to have the privacy principles perspective. We have to acknowledge that often society is changed by technology and not by laws. Guidelines and recommendation for implementers have to be developed. We need to identify areas in which a state of the art exists that is better than the common solution, and then, to derive guidelines. A privacy management system should be taken into consideration.

Second conversation: Privacy engineering process - a top down versus bottom up approach?

Top down: the Privacy Engineering & Assurance (PEAP) model

The privacy engineering process has to become an integral part of product development. Privacy requirements have to be translated into product requirements, and they have to be complemented by privacy assurance. It is integral to product development life cycle and must be complemented by privacy assurance to verify that security and privacy controls work. Can privacy requirements translate into product requirements? How to create 'privacy testing' and privacy coding training? One aim is to provide evidence to do a root cause analyses of problems. Another desirable goal is to be able to map privacy principles (e.g., ISO 27000) to privacy requirements. If we know what privacy requirements are, they are just one more form of quality assurance that can be measured and checked. We lack a mechanism to drive businesses and developers to pursue this approach.

A mapping of standard operating procedures is needed, as well as documentation of this knowledge. Privacy by design has to be transposed manageably and systematically into the product development life cycle. The problem extends from modelling in the very beginning until deployment and maintenance. Also, accountability has to be in the focus and present all phases of development.

Transposing privacy by design into manageable and systematic privacy engineering and assurance process (PEAP) is a 2 step process: (1) incorporating the seven privacy by design steps into privacy

policies, governance structures, set of principles, roles and responsibilities, (2) establishing standard operating procedures.

Bottom up: developers adapting tools and systems to realise privacy engineering

Software developers already have the tools, skills and technology to enable privacy by design, e.g., aspect oriented programming, cross-cutting concerns, annotations, how to validate work with unit and integration tests, and then to automate user acceptance tests. This could be as simple as tracking privacy themed annotations during a build process to implementing an ontology of annotations that express contextual data protection at run time. The context could depend on the user, the role, the origin, the operation, the time, the data and so on, varying the outcome and protection appropriately. Common software frameworks already allow for this approach and for it to be customised and maintained depending on the business and the product. This is not new - today's engineers know how to validate and test their work, and it is indeed now a widespread requirement for engineers to be able to demonstrate an ability to unit test, mock and automate. So what privacy by design, from the bottom-up needs, are enthusiasts and pioneers that do this and promote best practices for personal data protection. The privacy engineer is the key in bridging the technology and legal worlds. The privacy engineer has the ability to explain to the legal teams why privacy measurements are needed and on the other hand to crystalise the legal requirements in technology design and implementation. Also while they have the possibility to persuade the management, they also need the support of management to do this job and treat it with the same respect as any other quality management aspect or non-functional-requirement - throughout the product and software development life-cycle from business case to end-of-life. Of course, it is likely that adhering to this will not be possible in every development environment - for instance the "Mom-and-pop" or "garage" developer can keep an eye on this, but unlikely to be able to implement it at full scale. However, for many companies where technology is their bread and butter, this can be done.

So, act now; don't wait; do what you as an engineer can knowing you already have the tools and techniques available to you. The bottom up approach can produce tutorials, patterns, use cases, and cookbooks to aid developers, i.e., practical blueprints for use in common problems during the development life-cycle. These should not be just for or by developers, but also the legal team, the privacy officer, and the business, technical and product managers. At the other end, those companies that are inherently engineering based - in the true sense of "engineering" - can go further and help develop the standards and validation processes and tools that may at some day lead to a kind of privacy by design certification, while also becoming day-to-day instruments in the non-engineering development shops and drag-and-drop style tools that are often used and relied on in today's ephemeral "app world". Finally, all this has to be done in a supporting environment. Governance and the executive mandate to do this have to be present and take part not only to keep regulators happy, but also to build and leverage individual's data protection needs in innovative and respectful ways.

When this cannot or is not done, regulators have a spectrum of enforcement. This is true today, and will be even in a privacy-by-design based future. Unfortunately it is not enough to hope for compliance as there are and probably always will be some "bad" actors, either through lack of

diligence, ignorance or in some unfortunate cases - intent. Ultimately, enforcement can go as far as shutting down a business process as long as a process violates the regulation, and this can become a make-or-break situation. So not taking on board the principles of privacy by design can have serious consequences. Governance is key.

Third conversation: Real life conditions

Building privacy into a product is analogous to security, and a privacy engineering process needs to be integrated. Analogous to security, it can be integrated in the product design and development. So what stops privacy from already being part of the product life cycle? There is a lack of reward for businesses to implement privacy, since it does not increase their revenue. Also, privacy concerns have been perceived to hamper the development of technology.

How to tackle these issues? Some companies try to be pioneers in the privacy and data protection sector and to set new standards. Policies are needed that make outdated behaviour unacceptable (e.g., mail-servers using plain text communication). Also, we must learn how to better explain why privacy is needed.

Where to go from here? The participants were invited to share the issues and activities that they would like IPEN to focus on. Among them was the wish for 'privacy cookbook', not only one for developers, but also one for business process designers. A resource list should be made available. Libraries like PINQ (Privacy Integrated Queries) should be set up. Patterns and guidelines for privacy by design should be created.

Besides technical wishes, there were some more general demands. There is still a lack of awareness in the public why privacy should be in their interest. Campaigning to raise awareness is one possible way to address this.

Existing projects that are aimed at protecting privacy (e.g., Tor) would be interested in improved funding, in particular when their development is difficult to justify as a business case. Security technics and privacy mechanisms that already exist could be combined with protocols or software that is widely used (e.g., IMAP via SSL). Long-term development to fix fundamental flaws of the internet or the creation of new privacy friendly basic technologies need support.

Enforcement of compliance could be used as another means to make privacy and data protection more relevant for businesses. User privacy policies that are readable by machines can be developed. DPAs can encourage providers of services to upgrade their standards (e.g., scan servers whether they are encrypted). A concrete policy wish included the implementation of a regulation that app developers are not allowed to forward data without consent.