Annual Privacy Forum 2020
22-23 October
ONLINE EVENT
BRINGING RESEARCH & POLICY TOGETHER
www.privacyforum.eu

ANNUAL PRIVACY FORUM 2020 NOTES

BACKGROUND

The value of personal data in the online world has significantly increased over the last years, as electronic products, services and processes have permeated every fold of everyday life. Limitations in the transparency, the functionality and interconnectivity of online and communication services increases the risk of having personal data processed out of control of any accountable person or organization or simply becoming exposed to all sorts of privacy threats.

The EU legal framework on personal data protection is key in an effort to better control the processing of personal data while ensuring an adequate level of protection. Even the best legislative efforts cannot keep up to speed with the pace of innovative technology and business models that challenge the way personal data is processed and privacy is protected across the EU and beyond; therefore, examining what is at stake and where threats thereto originate from becomes of paramount importance.

Against this background, ENISA, DG CONNECT and the Católica University of Portugal, Lisbon School of Law organized the Annual Privacy Forum (APF) 2020.

Due to the COVID-19 pandemic, this edition of APF took place as an online event on the 22nd and 23rd of October 2020.

Twelve papers were accepted for publication (a 20% acceptance rate) and presented during the conference, after a thorough peer-reviewing process, on the basis of significance, novelty, and scientific quality. The papers were organized across four thematic areas, were presented in four respective paper sessions and are published by Springer LNCS Proceedings[1].

This document presents, in brief, the key points made during the conference and relevant conclusions, complementing the published proceedings.

[1] https://doi.org/10.1007/978-3-030-55196-4

DAY 1 - THURSDAY, OCTOBER 22, 2020

OPENING REMARKS

Andreas Mitrakas, Head of Data Security and Standardization Unit, welcomed the participants and opened the conference on behalf of ENISA's Executive Director, Mr. Juhan Lepassaar. The General Data Protection Regulation (GDPR) lays our security as a principle and as an obligation for all entities that process Personal Data. Cyber Security Strategies that have been launched from the beginning of the past decade, evolved to the NIS Directive and have supported a better security posture in several sectors. The Cybersecurity Act Regulation (CSA) laid the ground for the EU cybersecurity certification framework, towards increasing trust in ICT products, services and processes. Lastly, ENISA is also tasked to support the European Data Protection Board (EDPB), upon request, and provide advice, guidance and knowhow in relation to cybersecurity in data protection and privacy. During the previous years, ENISA has already worked around the concept of balancing of cyber security and data protection and privacy, providing many publications and reaching out to a broader public through the Annual Privacy Forum (APF), a collaborative event. The last editions have been organized back to back with the IPEN workshop, organized by the EDPS.

In security and privacy, interdependencies are multifaceted and significant, so a balancing is a recurrent endeavour. Data breaches are a reminder of the business risks and the data subjects risks. In view of the pandemic, the use of tracing applications to fight the disease puts public policy into new testing. The GDPR aspires to have a global impact and the Cybersecurity Act Regulation is likely to have a similar dimension, as we are moving gradually from a sovereign layer of control to a shared EU level of control, in relation to cybersecurity certification. ENISA, from its earlier years, has contributed to a clearer understanding of the risk-based approach in terms of determining security measures, an approach which is now the dominant one. In the area of certification, action is needed in relation to security measures, especially since some application areas are in need of such schemes. Art. 42 of the GDPR is hoped to provide solutions. New challenges will still be need to be taken into account especially for the sectors of supply chain, cloud computing and 5G.

PANEL SESSION I: DATA PSEUDONYMISATION: FROM THEORY TO PRACTICE

Panelists: Paolo Balboni (ECPC), Monir Azraoui (CNIL), Damien Desfontaines (Google) **Moderator**: Athena Bourka (ENISA)

Athena Bourka, introduced the panellists and the topic of the session. Pseudonymisation, a known de-identification technique gained additional tension with the GDPR as it can be seen both as a security measure and as a privacy by design technique. ENISA has provided specialized guidance during the previous three years.

Monir Azraoui presented the regulatory perspective of pseudonymisation, which is defined in art 4 (5) of the GDPR. Although data controllers believe that pseudonymised data is anonymised data, the GDPR clearly states that they are still personal data. Pseudonymisation can reduce the risks and can relax, to some extent, some legal protections, e.g. in the case of data breaches. He presented good and bad practices, from CNIL's audits. For the bad practices, he mentioned that: hashing cannot be though as an anonymization mechanism, simply removing direct identifiers is not anonymization, the use of bad cryptographic keys or obsolete hashing algorithms is insecure, and measures are needed to protect "additional information" like the cryptographic keys. For good practices he mentioned: including re-identification into risk-analysis, good protection of additional information, combining pseudonymisation with other security techniques - especially anonymization techniques. Mr. Azraoui presented practical examples of good techniques. In the first example, the data are partitioned into shares that are not related. The second example comes from the French Health Insurance Information System, where different databases corresponding to different regimes are combined into a single database through the use of two levels of pseudonymisation. The third example is the ticketing data in the public transportation system, where the secret key is split to different persons.

Paolo Balboni “provoked” the audience, with his presentation of Pseudonymisation as the ‘new normal’. Authorities (Art. 29, EDPS, AEPD, ENISA) have provided guidance with several opinions and publications. Since the concept of personal data under EU law is broad, full anonymization is difficult, so in most cases, data processed should be considered pseudonymised. The post-Schems situation might boost the use of such techniques, since in most cases adequate supplementary measures will be required, to protect personal data from risks arising from domestic provisions, in the third country, that provide access to foreign authorities without sufficient guarantees. Initial guidance by authorities indicate that encryption, anonymization or pseudonymisation might be candidates for such measures, although guidance from the EDPB is expected to be published soon. The Commission is also active, working with authorities, for the modernization of Standard Contractual Clauses and to discuss with US authorities a stronger data transfer framework. Other examples of sections where pseudonymisation is traditionally used are clinical trials, but also the Ad-Tech, where it can play a very important role. Mr. Balboni presented a Corporate Social Responsibility Framework developed at Maastricht University, made of 5 principles and 15 rules. The very first one is about Data Security by Design and includes pseudonymisation as one of the recommended techniques.

Damien Desfontaines presented Google’s differential privacy proposal, which provides an anonymisation solution. He introduced the basic concepts of the technique. With differential privacy, individual data cannot be inferred from a statistical set of results, since noise is added to these results. Privacy is respected, data can still be useful for the statistical purposes, but there is trade off in accuracy of the results. The method provides formal guarantees as it shows resistance to post processing and auxiliary knowledge, but, on the other hand one should cope with challenges like noise calibration to achieve a certain accuracy level, subtle implementation details and it is not suitable for small sets of data. The work of Google’s team is provided as open source . Google’s Covid-19 Community Mobility Reports make use of the technique. The reports use aggregated, anonymized data to chart movement trends over time by geography, across different high-level categories of places, comparing the data against a baseline. Desfontaines presented an example of how differential privacy was applied for this reports.

In the Q&A section of the panel session, Paolo Balboni stressed that the law places the responsibility for the selection of the proper techniques, including pseudonymisation, to the data controllers and not the regulators. Since more explicit indications are currently available, both from Supervisory Authorities and Courts, he expressed the hope and view that now is the right moment for organizations to start considering the operationalisation of pseudonymisation in their workflows. Monir Azraoui added that the CNIL is not a solution provider but provides advice. There is not a one fits all pseudonymiszation solution and the appropriate technique depends on the specific processing. Damien Desfontaines elaborated on how to find the right balance when adding noise in differential privacy, explaining that this choice is case specific and is depended on the level of accuracy and the level of formal guarantees that you get. He argued that the industry should be transparent when using differential privacy parameters, since that would provide feedback and help other engineers.

Desfontaines denied that by pausing location history, a data subject is denied access to services, but stressed that some services depend on location history to function properly. Paolo Balboni added that one should take also into account the contractual obligations between the parties, since by applying encryption or pseudonymisation techniques you cannot always provide the same level of service. In that respect, it is very important to carefully assess SLAs. Balboni invited everybody to think multidimensionally, since not everything is Data Protection and Cybersecurity.

Monir Azraoui mentioned that CNIL in the past, has seen anonymisation in event flow applications, with several techniques used, including hashing, generalization and randomization. The authority is also looking at many other techniques, depending on the specific application. For the case of Health DataHub, where data will be collected for scientific and research purposes, he clearly stated that according to CNIL the data are clearly pseudonymised, and not anonymised. Damien Desfontaines stressed the need for more open source tools, and that a list of tools for pseudonymisation would be beneficial. Paolo Balboni agreed that open source tools should be promoted and stated that the answer to the need many companies have for off-the-self tools is similar to that of the security measures, where a risk based approach is needed and there is no one-fits-all solution. The papers of ENISA present an overview of these techniques. Monir Azraoui agreed also, stating that pseudonymisation should first be selected as an appropriate risk mitigating solution and then the actual method can be found.

In the final minutes of the session, the panelists were asked for a one-minute advice. Damien Desfontaines stated that anonymisation is hard but feasible. Paolo Balboni argued that data protection and functionality should be both preserved. Monir Azraoui stated that anonymization should not be the holy grail of data controllers, but they should also consider pseudonymisation. The GDPR fosters trustworthy innovation.

DAY 2 - FRIDAY, OCTOBER 23, 2020

OPENING REMARKS

Jorge Pereira da Silva, Dean of School of Law of the Católica University of Portugal, greeted the participants and opened the second day of the conference. Mr. da Silva argued that we are living a redefining moment in the history of human rights. A new generation of rights, born or reborn in the last decade, are the result of a struggle to respond to technological threats. Interdisciplinary approaches are crucial to secure fundamental rights in these fields, so we need to build bridges between law, ethics, data and computer science and public policies. Jorge Pereira da Silva concluded his remarks with a sentence from philosopher José Ortega y Gasset on how technology influences and changes everything.

THE NIST PRIVACY FRAMEWORK

Naomi Lefkovitz, Senior Privacy Policy Advisor, argued on the added value of NIST's voluntary Privacy Framework, presented earlier this year. The overall objective of the framework is to help build trust, support ethical decision making and fulfilling compliance obligations, irrespective of the particular jurisdiction and facilitate communication and transparency. She explained, through a Venn diagram, that the notion of risk is similar but different, between Cybersecurity and Privacy. The understanding of risk is more mature for Cybersecurity and the diagram helps organizations realize that Privacy risks are connected to the various problems that individuals may encounter, due to the processing of their data. Privacy risks may result to organizational risks, since the direct impact on an individual might lead to loss of trust and customer abandonment, noncompliance costs, harm to reputation etc.

The structure of the Privacy Framework is modelled after NIST's Cybersecurity framework. The core provides an increasingly granular set of activities and outcomes. Profiles help organization prioritize and manage privacy risk. Implementation tiers are generalized benchmarks that organizations can use. Ms. Lefkovitz mentioned that, as a report from IAPP has shown, most privacy experts are using some form of the framework, probably due to its flexibility. She added that it provides a very easy way to communicate privacy to upper management and that it is easy for a company to use existing profiles from the Cybersecurity framework and through a gap analysis, update their requirements and control. NIST has developed a resource repository containing mappings between standards and regulations to the framework as well as privacy guidelines and tools. Work is ongoing with the goal to advance the framework and provide more guidance and tools.

Ms. Lefkovitz, noted, answering relevant questions, that the IAPP report indicates that the framework is mostly in use in the Health and IT sectors, but also in a variety of companies and that NIST welcomes contributions to the framework and the resource repository. She also stated that NIST is very much looking in AI and some parts of the framework are designed to address the issues on algorithmic bias. Finally, she argued that the framework is reconciling risk-based (US) and principle-based (EU) approaches, since the analysis is framed around 'problems' which covers both approaches.

WEB PRIVACY OR WEB (IN)SECURITY

Pedro Adão, associate professor at the University of Lisbon, argued that privacy may not always be achieved, since some times the technology itself may be the source of the issues. He specifically mentioned the case of cross-site leaks (XS-Leaks). In these attacks, the attack vector consists in exploring side-channels on the web/browsers such as timing, errors and size of request responses. This class of vulnerabilities is quite recent and browsers are now introducing techniques to deal with these attacks. In these attacks, the victim visits an attacker controlled web page and his private information is extracted by abusing regular endpoints of applications. In most of these attacks, the attacked servers are operating properly. XS-Leaks could become one of the top privacy threats in the near future.

Answering to questions, Pedro Adão stated that all browsers face the same challenges, because the attackers are exploiting the way the web is (or better was) designed. Vendors are currently working together trying to propose new web platform security features to solve this issue but there is no default solution that mitigates the entire class of vulnerabilities, and the ones that exist are not widely adopted. It is not an issue of the servers, since servers have (or better had) no knowledge on where the malicious request is coming from. Pedro Adão concluded the problem sits in "no man's land", but browser vendors are exploring solutions that provide applications with exact information about those requests (see fetch metadata). For the moment, users just have to wait for these features to be widely adopted.

PANEL SESSION II: TO TRACK AND TO GET TRACKED: NEW INNOVATIVE METHODS AND ADVANCEMENTS

Panelists: Marit Hansen (ULD), Rob van Eijk (Future of Privacy Forum), Fernando Silva (Banco de Portugal)
Moderator: Prokopios Drogkaris (ENISA)

Prokopios Drogkaris introduced the panelists and the topic of the panel. Tracking has become an indispensable part of our lifes and in many cases users are requesting for such a service. Tracking applications have been recently used to fight the pandemic.

Marit Hansen presented the experience of the supervisory authorities on tracking. She defined tracking as the act or process of following something or someone. It is widely used on the Internet, as the basis for targeted advertising. Tracing is different from tracking, since in a tracing application, the user is not followed, but provides traces. Hansen presented several forms of tracking, like cross-device tracking via ultrasound, fitness, WiFi or MAC address tracking, explaining that it is not always necessary to identify persons. The applicable data protection provisions are art. 3 (2) of the GDPR and the non-yet GDPR harmonized ePrivacy legislation. The ECJ, with its decision on the Planet49 case, affirmed that consent is required for tracking through cookies that are not strictly necessary for the provision of a service. In the real world, cookie consent seems to be a pain. However, consent is required for purposes like analytics and performance, social media, targeting and advertisements. In many websites, cookie consent banners are appearing, but with small usability. Targeting can also happen from observed or inferred data, a remark made by the EDPB in the recent guidelines 8/2020[2].

Rob van Eijk presented an example of how modern webpages are presenting ads, by using several small elements each of which might use its own tracking technology. A comparison of major browsers shows that they behave differently in connection to third-party cookies. Google Chrome, the dominant browser, will join its competitors and block these cookies, but not before 2022. In the post-DNT period, new security and privacy by design techniques are used in browsers (SameSite cookies, Strict site isolation, FLoC),

[2] https://edpb.europa.eu/our-work-tools/public-consultations-art-704/2020/guidelines-082020-targeting-social-media-users_en

transparency frameworks are developed and mandatorily used (like in iOS 14) and more regulatory requirements (e.g. CCPA, CPRA) enforce global privacy controls, like the "Do not sell" initiative.

Fernando Silva presented on new and innovative tracking methods, closer to the state of the art of tracking and tracing. Tracking is used to create heat maps used for several purposes, including advertising, through scanning MAC addresses and SSIDs. Bluetooth's MAC address is exploited in the same way and provides more accuracy. It has even been observed that following OS updates, a deactivated Bluetooth can be reactived. Intelligent video analytics use visual intelligence to provide video metadata but can be associated with facial recognition systems in order to track persons. Fernando Silva mentioned also RFIDs, GPS, mobile data, audio beacons and several other tracking and tracing techniques. He presented articles about facial recognition in China to track minorities and argued that specific applications collect vast amount of surveillance data, which is a risk by itself, even if the original purpose is not surveillance.

The use of masks during the pandemic affects the accuracy of facial recognition. But contact tracing became easily accepted, whereas that wouldn't be possible one year ago. The partnership between Google and Apple creates mistrust. Experts fear that tracing could become the 'new normal' even after the pandemic. Risks are already present. In some cases it was observed that contact tracing data were misused by authorities. In the US, a company has rolled out employee covid-19 tracking that can be used as covert surveillance. COVID tracing apps in Europe exhibit vulnerabilities into protecting the Bluetooth rolling proximity identifier, which is key to safeguarding privacy. For contact tracing to work, people need to trust companies and governments, while their track record is not perfect. Silva noted "Trust but verify!". Concluding, he stated that the main privacy risks arise from the pervasive use of technologies that were meant for another purpose. Transparency is crucial. According to Fernando Silva, contact tracing apps have no use at all, but only serve to open Pandora 's Box.

In the Q&A section of that panel session the first question was on configurability. Fernando Silva pointed out that it cannot always the answer, e.g. in the case of employees in their workplace. Privacy by default is not respected when there's imbalance of powers. Rob van Eijk noted that in the browser domain, most solutions use chromium and webkit. He added that for configurability to be effective, privacy by default should be respected, since many persons, especially minorities, lack the luxury to change the options. So, it is imperative that services process, by default, only the amount of data that are necessary for the service requested by the user. He concluded that although the GDPR provides rights, the root of the problem is the use of stable identifiers. Marit Hansen agreed that consent might not be the right solution, not only because it is difficult for the date subject, but also because it is not ideal for the data controller. She agreed that the by default principle should be used even more. She also noted that tracking is now changing, but that's probably because controllers are forced to change, due to the court rulings. But this change of direction is not always usable and sometimes controllers blame the regulators for that. She noted that even now, dark patterns are used, which is an ethical and also legal problem, while also mentioning the Deceived by design report[3]. A possible solution would be some sort of privacy hardening. Removing all not necessary components and offer a basic service to the users. These areas are not explicitly laid down in laws, they could be interpreted as part of data protection by default, but work needs to be done, since 'greedy' data companies might seem to adhere to the law but always but moving on the edge of lawfulness.

In their closure remarks, Marit Hansen stated that she'd liked to have guarantees that there are tracking free spaces. Rob van Eijk is optimistic about contextual advertising and innovation on that field, since research shows that visitors interact longer with the content. Fernando Silva praised the significance of transparency from application providers and asked participants to think a bit before "clicking".

CLOSING REMARKS

Andreas Mitrakas thanked the participants and gave the floor to Nils Gruschka, Associate Professor at the University of Oslo, who announced the next edition of the Annual Privacy Forum. APF 2021 is going to take place in Oslo, on 17 and 18 of June 2021 and will be co-organized by the University of Oslo.

[3] https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf

CONCLUSIONS

The conference covered important aspects of privacy and cybersecurity, engaging more than 500 participants into a fruitful exchange of views. Some of the main findings and/or open issues are as follows:

- **Enhancement of the principle of accountability in relation to AI**. Accountability is the best answer to the risks of AI. That can be achieved by making sure that those who create AI applications are responsible for their creations and ensure the perfect balance, even if the system is autonomous enough to take its own decisions.

- **Privacy and data protection do not stand in the way of innovation**. The GDPR fosters trustworthy innovation. Manufacturers and controllers should provide solutions that respect privacy while meeting their goals. Legal developments might boost innovation in fields like contextual advertising, since research shows that visitors interact longer with the content.

- **Data protection and functionality can both be preserved**. One should take also into account the contractual obligations between the involved parties. Encryption or pseudonymisation are not mandatory techniques but should be examined within the context of each specific case.

- **Open and transparent resources and guidance is necessary for privacy techniques**. Organizations should consider the proper techniques, with or without guidance from regulators. Open source tools and should be promoted although off-the-self tools should be used with a risk based approach. The industry should be transparent when using parameters in their solutions, since that would provide feedback and help other engineers. Regulators are not solution providers but can provide advice.

- **Transparency is the answer to many problems**. Application providers need to provide more transparency for their solutions. Standardized purposes can minimize legal uncertainty by offering more transparency.

- **Privacy by default should be respected, especially in cookies and fingerprinting**. Tracking technologies are changing, due to court rulings, but cookie consent banners that are appearing, lack usability and in several cases, employ dark patterns. Controllers should be able to provide guarantees that there are tracking-free spaces.