e-SIDES Ethical and Societal Implications of Data Sciences

Assessment of Privacy-Preserving Big Data Technologies Under Development

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731873

e-SIDES

Ethical and Societal Implications of Data Sciences

About the e-SIDES project

Data-driven innovation is deeply transforming society and the economy. Although there are potentially enormous economic and social benefits, this innovation also brings new challenges for individual and collective privacy, security, as well as democracy and participation. The main objective of the CSA e-SIDES is to complement the research on privacy-preserving big data technologies, by analysing, mapping and clearly identifying the main societal and ethical challenges emerging from the adoption of big data technologies, conforming to the principles of responsible research and innovation; setting up and organising a sustainable dialogue between industry, research and social actors, as well as networking with the main Research and Innovation Actions and Large Scale Pilots and other framework program projects interested in these issues. It will investigate stakeholders' concerns, and collect their input, framing these results in a clear conceptual framework showing the potential trade-offs between conflicting needs and providing a basis to validate privacy-preserving technologies. It will prepare and widely disseminate community shared conclusions and recommendations highlighting the best way to ultimately build confidence of citizens and businesses towards big data and the data economy.

Deliverable D5.1. Assessment of Existing Technologies Under Development

Find more at: https://e-sides.eu/assets/media/deliverable-d5.1-assessment-of-existing-technologies-under-development-1.0.pdf

About this white paper

This white paper is based on Deliverable 5.1 of the e-SIDES project, which had the main objective to validate privacy-preserving big data technologies under development as well as their implementation.
The main research question under this deliverable was the following: *How are security and privacy features in big data-driven innovation projects—which we call privacy-preserving technologies—suitable to deal with the ethical, legal and societal values that come under pressure in specific big data application contexts?* We also looked at how the ethical, legal and societal implications of such technologies are measured.

IDC ANALYZE THE FUTURE

eLAW LEIDEN

Fraunhofer

Grant Agreement number: 731873

Methods

With the aim to perform an assessment of existing privacy-preserving technologies under development we first identified a set of data protection, social impact assessment, and responsible research and innovation (RRI) assessment tools:

- Art. 29 WP Guidelines on DPIA,
- The United Nations Global Pool Data Protection Impact Assessment,
- Data Ethics Impact Assessment,
- Data Ethics Canvas of the Open Data Institute, Software Impact Assessment;
- ISO/IEC 29134:2017 Information technology -- Security techniques -- Guidelines for privacy impact assessment, responsible innovation assessment tools
- Applied Ethics Toolkit developed through Deliverable 2.2 of e-SIDES.

Secondly, we identified three big data areas in order to assess how are values and the implications of privacy-preserving technologies and big data solutions assessed. These three areas were the following:

- Healthcare;
- Transportation and smart cities;
- Web browsing and third-party tracking.

Each of these areas was chosen for a set of reasons. Healthcare was chosen due to the shrinking prices of medical examinations, the growing amount of wearable mobile devices, the increasing ageing population, and because of the ever-growing demand for data and analytics has become a common phenomenon. Transportation and smart cities were chosen because the accumulation of purposes for which big data is used in smart cities has grown in recent years and looks as though it will continue on this trajectory. Web-surfing and third-party tracking was selected because big data developments in this sector have major implications.

We used desk research and questionnaires to interview the identified Horizon 2020 ICT projects in the three areas mentioned above.

HEALTHCARE

We assessed the following Horizon 2020 projects in this area:

- BODYPASS (https://www.bodypass.eu/)
- MyHealthMyData (http://www.myhealthmydata.eu/)
- CLARUS (http://clarussecure.eu/)
- BigMedilytics (https://www.bigmedilytics.eu/)

Under Article 9 of the EU's General Data Protection Regulation (GDPR), 'the lawfulness of processing special categories of data', healthcare information counts as a special category of personal data and requires advanced protection. There is a significant amount of research assessing the implications of privacy-preserving and security technologies for healthcare. Academics on healthcare data[1] have demonstrated that assessing these impacts can be context-dependent and device-dependent as the ways in which data is collected and processed across technological tools differs per healthcare application. For instance, a widely used privacy-preserving technology in healthcare is homomorphic encryption. Certain researchers argue that complex algorithms must be disintegrated into simpler operations; a framework using homomorphic encryption, for instance, must be disintegrated in terms of simple homomorphic additions and multiplications which can also be considered as very useful safeguards for successful societal, ethical, and legal assessments.

We concluded that based upon insights into the ethical, legal, and social impacts assessment tools and the BodyPass, MyHealthMy Data, CLARUS and BigMedilytics projects, the healthcare sector is scattered. Despite the strictest data protection rules apply for healthcare data, diverse privacy-preserving technologies and types are implemented for epidemiological, diagnostic and other reasons. Each of the four projects in light of our findings could benefit from relying upon one or more of the impact assessment tools we identified before.

TRANSPORTATION AND SMART CITIES

Transportation and smart cities are characterised by increasing ubiquity and combinability of digital and autonomous technologies. They allow for the exponential lengthening of the big data lifecycle. Within this context we assessed the following Horizon 2020 projects in this area:

- SPECIAL (https://www.specialprivacy.eu/)
- AEGIS (https://www.aegis-bigdata.eu/)
- Transforming Transport (https://transformingtransport.eu/)

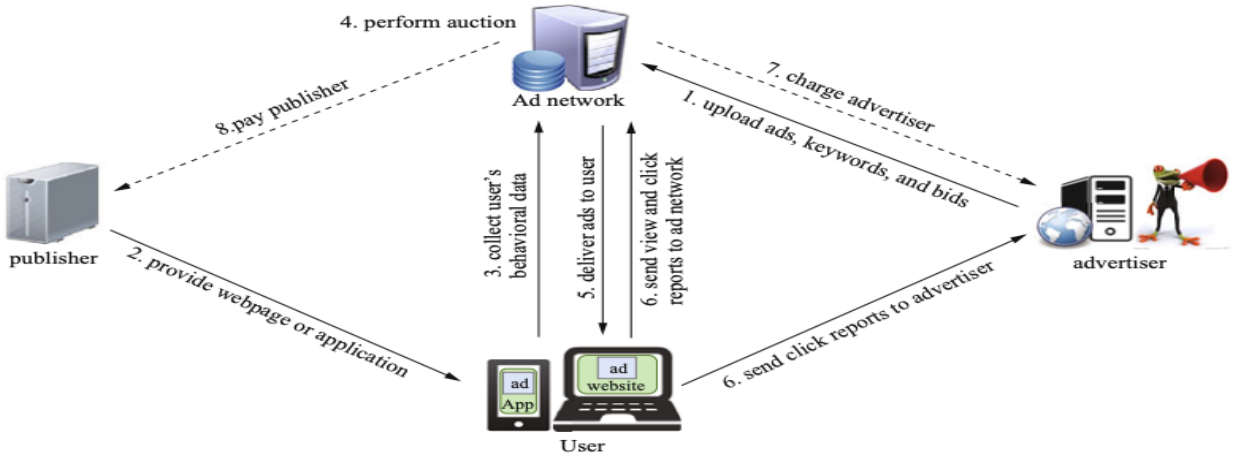
With respect to projects focusing on transportation and smart cities we made the following observations:

- Projects are aware of the impacts of the big data solutions they develop on the society. They assess these impacts but usually not in a highly systematic way or following a rigorous methodology. Some established ethics advisory boards.
- Projects tend to integrate common privacy-preserving technologies into their solutions. If possible, personal data is anonymised or at least processed, stored and transmitted in ways that are secure, for instance, by using encryption. Policy enforcement and user control also receive considerable attention.
- Projects generally stress that they are compliant with applicable laws. The GDPR, its predecessor, or national privacy laws are frequently mentioned. Applicable corporate policies are usually taken into account by formulating system requirements accordingly.
- The extent to which projects pay attention to societal and ethical issues varies greatly. If projects take measures to assess impacts and mitigate them, they usually implement preventive measures. Reactive measures do not seem to play a significant role.
- The alignment of people, processes, and technology is typically not made explicit. However, project documentation clearly shows that projects acknowledge that people and processes are at least as important as technology.
- The adoption of best practices is very common among projects. Projects build upon lessons learned and implement whatever has proven useful and does not interfere with the project's specific objectives.
- Projects focusing on transportation and smart cities do not generally see a conflict between protecting economic interests and avoiding undesired ethical and societal implications. On the one hand, economic interests are not overvalued and, on the other hand, undesired ethical and societal implications than cannot be mitigated are considered as show-stoppers.
- The measures taken in the transportation and smart city context do not seem to differ much from those in the application contexts. Personal data is not essential for many big data scenarios related to transportation and smart cities.

WEB-SURFING AND THIRD-PARTY TRACKING

We looked beyond the Horizon 2020 projects and assessed this area more broadly because innovation in this area touches different sectors at the same time.

This context permeates a wide variety of uses of the Internet. As many websites and applications are funded through the sale of advertisements, technologies have grown to make these advertisements more effective. A common way in which this is currently performed is through the use of third-party tracking, where third-party cookies are used to monitor user behaviour across sites, usually with the intention to present targeted or more relevant ads to the user.



We concluded, while also including the Horizon 2020 Papaya project into our assessment, that a wide variety of different privacy-preserving technologies are under development in the web browsing and third-party tracking context. Some of the technologies can be employed unilaterally by one of the relevant parties (ad-networks, users, and to a lesser extent publishers), while other solutions require cooperation between these parties. Those technologies that require the participation of ad-networks have not seen much utilisation perhaps due to costs and lost revenue, while unilateral technologies employed by users to protect their privacy are actively developed and used. As this context develops, cooperative solutions that protect the privacy of individuals, while still being able to provide personalised features and advertisements, offer the most beneficial way forward.

CONCLUSIONS

From assessing the listed ICT-14, ICT-15 and ICT-16 projects and the development and level of implementation of privacy-preserving technologies, we have drawn the following general conclusions:

- The emergence of big data changes the protection of privacy as well as the relevance of other related issues, such as confidentiality and utility, significantly;
- The evaluation of projects also demonstrated that by moving towards privacy, the granularity of data for business purposes becomes more difficult to exploit;
- The rapid innovative developments and the variety of ways in which privacy-preserving technologies are used, even in combination with each other, e.g.: in healthcare, also underline the importance of time during the assessment of impacts;
- In all three contexts privacy-preserving technologies and data sharing methods serve to optimise processes that are increasingly geared toward improving predictions on the future behaviour of citizens;
- Beyond the continuity in assessing the present impact of technologies, impact assessments should also embrace considerations regarding potential impacts in the future.

To know more about e-SIDES:

www.e-sides.eu

To contact us:

info@e-sides.eu

@eSIDES_eu

Grant Agreement number: 731873