RESEARCH OF METHODS AND MEANS OF ORGANIZATION DATA PROTECTION BASED ON THE ANALYSIS OF CIRCULATING INFORMATION

Sagatov Miraziz Varisovich
Doctor of technical sciences, head of the "Information Technologies" department,Tashkent University of Information Technologiesnamed after Muhammad al-Khwarizmi, Uzbekistan
Nomozov Mansurbek Nurali o'g'li
Master's degree, Faculty of Cyber-Security,Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

https://doi.org/10.5281/zenodo.7445130

Abstract. *Technological advancements over the last years have impacted the way our personal data is being shared and processed. The evolution of technology has brought forward new techniques to share, process and store data. This has generated new models of data (including personal data) processing, but also introduced new threats and difficulties for the end user to understand and control the processing. Continuous online presence of end users has resulted in an increased processing of large amounts of personal data at daily basis. Think of online shopping or using a mobile application to navigate to a specific location or contact friends and family. The whole data lifecycle has been augmented with many actors being involved and eventually end users not being able to fully understand and control who, for how long and for what purpose has access to their personal data.*

Keywords*: Information Security, Risk Assessment, Security Framework, Data Protection, GDPR.*

ИССЛЕДОВАНИЕ МЕТОДОВ И СРЕДСТВ ЗАЩИТЫ ДАННЫХ ОРГАНИЗАЦИЙ НА ОСНОВЕ АНАЛИЗА ОБОРОТНОЙ ИНФОРМАЦИИ

Аннотация*. Технологические достижения последних лет повлияли на то, как наши личные данные передаются и обрабатываются. Эволюция технологий привела к появлению новых методов обмена, обработки и хранения данных. Это породило новые модели обработки данных (включая персональные данные), но также создало новые угрозы и трудности для конечного пользователя, чтобы понять и контролировать обработку. Постоянное присутствие конечных пользователей в Интернете привело к увеличению ежедневной обработки больших объемов персональных данных. Подумайте о покупках в Интернете или использовании мобильного приложения для навигации в определенном месте или для связи с друзьями и семьей. Весь жизненный цикл данных был расширен за счет участия многих участников, и в конечном итоге конечные пользователи не могут полностью понять и контролировать, кто, как долго и с какой целью имеет доступ к их личным данным.*

Ключевые слова: *Информационная безопасность, оценка рисков, структура безопасности, защита данных, GDPR.*

Introduction. These new technologies have often been introduced without a prior assessment of the impact on privacy and data protection. In this context, processing of personal data is often characterised by the absence of a predetermined purpose and by the discovery of new correlations between the observed phenomena, for example in the case of big data or machine learning. This modus operandi conflicts essentially with the principles of necessity and purpose limitation, as these are stipulated by the GDPR. Blockchain and distributed ledger

technologies, as another example, offer the opportunity of replacing intermediation-based transactions, but at the potential expense of a substantial loss of individuals' control over their data, which remain visible in the chain by all blockchain participants, as long as it is active or perhaps even beyond that. This, depending of course on the use case, contradicts the GDPR principle of data minimization, and constitutes a severe obstacle for the exercise of the right to deletion by data subjects. Lastly, Artificial Intelligence systems might be empowered to take decisions with some degree of autonomy to achieve specific goals, for example in credit score evaluation in the finance domain. Such autonomy might very well be in conflict with the prerequisites of human agency over machines and self-determination, both at the heart of personal data protection and the GDPR.

Materials. A key element in any data protection concept is the enablement of human individuals to exercise their data protection rights themselves. This involves both access to information on data processing (transparency) and the ability to influence processing of their personal information within the realm of a data controller or data processor (intervenability). In this respect, a multitude of approaches and topics emerged from the privacy research community that can help implementing these rights and correlated services at data processing institutions. In this chapter, we present a selection of the most relevant ones

Methods. Similar to the right of access, the other data subjects' rights to erasure, rectification, blocking, restriction of processing, etc. can also be implemented in an equivalent way as dedicated services. Here, the infrastructure for the right of access services turns out to be very helpful, as it allows to easily identify all the data stores affected by the particular request. Also, a notification that a demand for the right to erasure or rectification was triggered can be sent to the data processors (or data controllers) responsible for this data.

Results. Private information retrieval (PIR) is a cryptographic technique which allows a user to recover an entry in a database without revealing to the data custodian (e.g. the database owner or administrator) which element has been queried [38]. This is why it can be used as a data minimization technique by data controllers. Let's assume that a company wants to provide access to a database to its customers. In a default setting, each time a customer makes an access to the database, the data custodian knows which entry has been accessed. Over time, the data controller will be able to identify which database entries are of interest to the customers. By implementing private information retrieval, the data controller minimizes the amount of information revealed on what was accessed PIR prevents the data controller from learning which entries have been accessed. Private information retrieval allows a user to recover an entry in a database without which element was queried. Private Information Retrieval Models There are two main models of private information retrieval. The first model is Computational Private Information Retrieval and there is only one server storing the database. This model is considered to provide better level of protection but has limitations with regards to the connections that can be established to the server and the database. In the second model, Information Theoretic Private Information Retrieval, the database is stored on several servers which are controlled by different owners. This model allows for better communication complexity but it is assumed that the servers do not collude or exchange information. Additional information on PIR are available.

Having such an automated right of access service as part of an organization's internal or external management systems reduces greatly manual efforts when it comes to excessive amounts of right of access requests. While employees can easily come to their limits when

demand grows up, technical infrastructure is usually easier to scale. Depending on the amount of such requests, such automated system may deliver significant savings to the organization. At the same time connecting to the right of access service each new data storage, data sink or an additional data processor that gets an individual's data may also improve data management capabilities of the organization as a whole. Requests concerning data locations, data forwards, business partners involved in data processing, etc. can all be answered rather straight-forwardly just from the existing data flow infrastructures created and maintained for the right of access service. On the downside, implementing such service requires additional processes to be defined, developed, and deployed along with the core functional services for data processing. This brings up a set of additional challenges to consider:

• Authorization: A human individual is only allowed to see and investigate its own personal data, not that of other data subjects. Hence, there must be some (technical Replying to a right of access request can be challenging in large, complex data processing networks with a multitude of data processors) means of authentication in place, to verify the authorization of the demanding individual. This authorization must, of course, guarantee validity of the authorization, hence must rely on high-level security techniques to validate the identity of a human individual. This may involve e.g. passport validation, two-factor authentication, or other similar means.

• Delegated Authorization: Sometimes, delegation of the right of access is possible, e.g. for under age children, legal custodians, attorneys, etc. In such cases, the authorization of the right of access request must be validated not just by verifying the identity of the demanding individual, but also by means of verifying the legal grounds for the transfer of authorization. Depending on the type of delegation of rights, this task may become arbitrarily complex (see also below).

• Risk of Data Breach: Disclosing the full set of personal data of one data subject to another data subject without valid authorization is equivalent to a severe data breach, which itself manifests a violation of the GDPR. At the same time, there may be a substantial interest in such right of access services by other actors than the concerned data subject, such as hackers, media, law enforcement, or relatives. Hence, the security risk for operating such a service is not negligible.

• Completeness: As was evident from the Schrems court cases, achieving completeness of the data disclosed in response to a right of access request is challenging. A recent study showed that only about 10% of companies provided a complete stack of customer data when demanded. However, an incomplete response to a right of access is a violation, and would hence render the utilization of such right of access service ineffective. The main challenge here is how to identify all the data that belongs to a certain data subject in the huge set of data stores typically found at large data controller or data processor organizations. Sometimes, this task maps to querying databases with the customer identifier of the data subject (if available), but it may also include skimming through vast amounts of archived data, file systems, backups, derived data sets, or other type of information that is no longer directly and easily accessible or properly linked to the data subject's customer identifier – if such identifier exists at all.

• Correctness: Similar to completeness, the data disclosed to the data subject must be correct, hence many not contain abbreviations, aggregations, internal censorship, or other access-blocking means. Also, its integrity must be maintained when delivered to the requesting individual. Hence, the implementation of such a right of access service must utilize sound

technical means to guarantee correctness and integrity of the data contained in the response to the right of access demand.

• Volume: Personal profiles of active data subjects typically grow in size over the time of utilization of a service. Hence, the size of the response to a right of access request can easily grow into huge amounts of data. This causes a technical challenge of delivering the data to the requesting individual by reasonable means. E.g. the maximum size of allowed e-mail attachment can be easily reached, rendering an information mail as response to a right of access request infeasible. Print-outs are not just environmentally problematic but also do not fulfil the common requirements concerning right of access responses nor the demand for data portability as define. Common solutions consist in web driven downloads of compressed data archives via HTTP(S) or FTP(S), which also have some technical challenges for low-bandwidth areas.

Conclusion: To conclude the thesis it makes a contribution to literature by presenting a new DPIA framework, which has been developed through the analysis of three risk assessment methods: OCTAVE Allegro, ISO/IEC 27005:2011 and NIST 800-30, the GDPR in combination with data collected from two companies. Currently, there is no standard framework for conducting DPIAs, and there are only guides on how to perform them. In the current situation, in which every company and organization within the EU or those outside the EU region which are processing personal data of the EU citizens, these companies and organizations must be compliant with the GDPR. Thus, it is important to standardize this process. A new standardized way of performing the DPIA will allow companies to be consistent in evaluating risks regarding the personal data and suggest concrete measure for the implementation of risk-treatment controls. Further, this thesis makes a contribution to practice and management by introducing a framework for identifying and treating only risks related to personal data. Concerning practice, it will speed up the process of identifying the risks to personal data, quickly analyze the gaps regarding the security measures inside the organization, identify solutions to address the gaps and provide a roadmap to reduce or avoid the probability of a data breach occurrence. In addition, pertaining to management, the framework facilitates to spend less effort, resources and money during assessments, by limiting the analysis only to the business processes which involve personal data. The implementation of the DPIA framework inside an organization will help improve awareness of the data protection risks, help manage the risks to the rights and freedoms of natural persons resulting from the processing of personal data by assessing them and determining the measures to address them and it will help in demonstrating compliance with the GDPR. The reputation of the organization will benefit from implementing such framework, because it will improve the trust and the confidence of the data subjects that their personal data is secured. It will also help organizations avoid the huge fines imposed by the GDPR, up to €20 million or 4% of the company's global annual turnover of the previous financial year, whichever is higher, by becoming compliant with the regulation and enhancing the security against data breaches. 44 The framework presented in this thesis could become a new standard for DPIAs. Procedures could be defined in order to help execute the risk assessment process effectively, document and manage risk remediation measures. The new standard would enable companies to comply with the GDPR and speed up the introduction and adoption of new business processes which involve personal data. Regarding future research, additional risk assessment frameworks should be analyzed. Further, a detailed comparison on different DPIA guides should be conducted in order to improve the DPIA framework proposed in this thesis. Further, suggested is

also that the Annexes of the ISO 27005 should be adapted in order to address personal data risks. For example, threats regarding the rights and freedoms of data subjects should be added. The outcome of this research, the DPIA framework, should be considered as a first draft of such a standardized DPIA process. With the volume of data processed around the globe growing rapidly and attackers, breaches, and fraud methodologies continuously evolving, the responsibility of companies and organizations to protect personal data has increased. The necessity of standardization of data protection worldwide has become obvious.

REFERENCES

1. Reason, P & Bradbury, H. (Eds.). (2018). Handbook of action research: Participative inquiry and practice, Sage Publication, London, UK.
2. Maung, K. Sein, Henfridsson, O. Purao S. Rossi, M., Lindgren R. (2021). Action Design Research, MIS Quarterly, 35(1).
3. Larman C. (2020). Agile and Iterative Development: A Manager's Guide, Addison Wesley Professional.
4. Hevner, A.R., March, S.T., and Park, J. (2019). Design Research in Information Systems Research. MIS Quarterly, 28(1), pp. 75-105.
5. Vaishnavi, V., Kuechler, W., and Petter, S. (2020). Design Science Research in Information Systems,
6. Peffers K. Tuunanen, T., Rothenberger M. Chatterjee S. (2017). A Design Science Research Methodology for Information Systems Research, 24(3), pp.45-78
7. Patton, M. Q. (2020). Qualitative evaluation and research methods, Sage Publications, London, UK.
8. Kvale, S. and Brinkmann, S. (2019) InterViews: learning the craft of qualitative research interviewing, Sage Publications, LA, USA.
9. Fontana, A. and Frey, J. (2019). Interviewing, in (eds) Denzin, N. and Lincoln, Y., Handbook of qualitative research, Sage Publications, Thousand Oaks, CA, USA.
10. Miles, M. and Huberman, M. (2020). An expanded sourcebook – Qualitative Data Analysis, 2nd ed, Sage Publications, Thousand Oaks, CA, USA.
11. Larry B. Christensen, R. Burke Johnson, Lisa A. Turner, (2018). Research Methods, Design and Analysis, 12th ed, Pearson Education Limited.
12. John W. Creswell, (2017) Research Design, Qualitative, Quantitative and Mixed Methods Approaches, 4th ed, Sage Publication, London, UK 46
13. Haes S., Debreceny R., Van Grembergen W. (2019). Understanding the Core Concepts in COBIT 5, Information Systems Audit and Control Association Journal, Vol 5, pp. 1-8
14. ISO/IEC 27005, International Standard, 15.06.2008
15. Bendtsen M. (2020). Risk analysis review