*Journal of Institute of Science and Technology 2019, **24(1)**: 1-5*
ISSN: 2469-9062 (print), 2467-9240 (e)
Doi: http://doi.org/10.3126/jist.v24i1.24620

Review Article

DATA PROTECTION IN THE EU AND ITS IMPLICATIONS ON SOFTWARE DEVELOPMENT OUTSIDE THE EU

Ralf Kneuper

Department of IT and Technology, IUBH University of Applied Sciences– Distance Learning, Kaiserplatz 1, 83435 Bad Reichenhall, Germany
Corresponding author: r.kneuper@iubh-fernstudium.de; www.iubh-fernstudium.de

*(**Received:** February 12, 2019; **Revised**: Mach 21, 2019; **Accepted:** March 29, 2019)*

ABSTRACT

In May 2018, the *General Data Protection Regulation* (GDPR 2016) came into effect in the European Union (EU), defining requirements on how to handle personal data of EU citizens. This report discusses the effects of this regulation on software development organisations outside the EU, and summarises the software requirements that result from GDPR and therefore apply to most information technology (IT) systems that will handle data of individuals based in the EU.

Keywords: Data protection, GDPR, Information technology, Software development

INTRODUCTION

In May 2018, the *General Data Protection Regulation* (GDPR 2016) came into effect in the European Union (EU), defining requirements on how to handle personal data of EU citizens. While at first sight, this only affects organisations within the EU, a closer look shows that the GDPR also applies to the processing of personal data performed in the context of goods and services provided to EU citizens, independent of where the processing is performed, and to the processing of personal data related to the monitoring of the behaviour of persons within the EU (Art. 3GDPR). Apart from these direct effects, the GDPR indirectly affects software development organisations worldwide whose customers have to conform to the GDPR. These organisations, wherever they themselves may be based, will have to address the GDPR requirements in their products since otherwise their customers may not be allowed to use them. The current paper therefore provides an overview of these requirements.

In most cases, there is some overlap between these requirements from EU legislation and data protection legislation in the region where the software development organisation is based. Since many of these organisations are based in Asia, the current paper will provide some pointers to such regulations in Asia as examples, without however going into details because that would go beyond the scope of this paper.

MATERIALS AND METHODS

This paper is mainly based on an analysis of the GDPR. Additionally, a short literature review was performed, with particular emphasis on those parts of the GDPR that affect software development and IT service provision. After giving a summary of the concepts of data protection and the GDPR, this report presented the main requirements of GDPR, with focus on those requirements that need to be addressed by software development, including a discussion of the functionality needed as a result.

DATA PROTECTION AND THE GDPR

This introduces the concept of data protection in general and in particular the European data protection legislation GDPR and its relevance to software development outside Europe, using Asia as an example.

Data protection

In spite of its name, data protection, also known as privacy, is not concerned with the protection of (confidential or sensitive) data as such, but with the protection of individuals from misuse of their personal data. For example, the design of a new product may be highly confidential from the point of view of the company that developed it, but in general is not covered by data protection. Data protection only refers to personal data as defined below.

The main goal of data protection is to give individuals control over their personal data, as one aspect of their human rights. This viewpoint may for example be found in the *Charter of fundamental rights of the European Union*: "Everyone has the right to the protection of personal data concerning him or her", with similar regulations stated in many other charters and constitutions, for example in Art. 28 of the Nepali *2007 Interim Constitution* (Greenleaf 2013).

In some cases, this may lead to a conflict with the right to *freedom of expression*, and different cultures come to different conclusions when balancing these two rights. As a result of this and other decisions about the relative importance of data protection, different countries sometimes have different requirements in their data

protection legislation, even though over the last few years, world-wide data protection legislation did converge to some extent.

For example, various Asian countries introduced data protection legislation such as the *Philippines Data Privacy Act* as updated in 2016, the *Singapore Personal Data Protection Act*, and the *Indian Data Protection Law* currently under discussion. Nepal's *Right to Information Act* of 2007 defines a number of similar data privacy regulations, despite its somewhat different focus (Greenleaf 2013). Though not legally binding, the Asia-Pacific Economic Cooperation's *APEC Privacy Framework* (APEC 2015) also states similar requirements. Of course, the requirements stated in these different laws are not identical, but in most cases are rather similar. China is a somewhat different case since here the "Social Credit System" increases surveillance of individuals by the state, while the *Data Privacy Standard* introduced in 2018 defines data protection requirements for private business similar to those of the GDPR (Magee 2018).

Overall, software development organisations world-wide will therefore have to adhere to their national data protection legislation as well as to the EU legislation discussed in this paper. It is important in this context to note the relationship and the difference between data protection and IT security. IT security starts from the viewpoint of the organisation and considers how to protect the organisation's own data. In data protection, the organisation needs to consider how to protect data about other people, often outside the organisation, which is one of the reasons why it is required by law and not left to the organisations involved. Data protection and IT security are both concerned with the protection of data but against different types of threats, even though the methods and tools used to do so will overlap to a considerable extent. Without adequate IT security, data protection cannot be implemented. The IT security requirements are outlined especially in Art. 32 GDPR and include the classic CIA triade (Confidentiality, Integrity and Availability) but also resilience as a requirement.

Personal data

As stated above, data protection only applies to personal data, defined as "any information relating to an identified or identifiable natural person" (Art. 4, item 1 GDPR). Examples of such personal data range from very simple data such as "X has the email address ...", via "the user with IP address X has visited the website Y", to rather critical and highly confidential data such as health data. These examples show that the level of protection needed may vary considerably, but the GDPR starts from the basic assumption that it should be up to the individuals concerned to decide about the amount of processing performed on their personal data.

To distinguish different risk levels, GDPR distinguishes two categories of personal data. Apart from normal personal data, there are "special categories of personal data" with additional restrictions on their processing, and additional requirements on their protection. Special categories of personal data involve "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, ... genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation"(Art. 9 GDPR). Nevertheless, organisations will have to go beyond these two categories of personal data to implement data protection and analyse in more detail the risks to the data subjects that arise from the processing of their data.

For some applications, it is possible to work with anonymous rather than personal data, in which case all the following statements on the processing of personal data are no longer applicable. However, when doing so one needs to verify that the "anonymized" data indeed no longer allow the identification of the individuals concerned. Just deleting the name or some other identifier in most cases is not sufficient to make the data genuinely anonymous. Similarly, using a tracking cookie on a website to identify repeated visits by the same individual defines a pseudonym but does not lead to anonymous data.

Roles involved in data protection

The GDPR distinguishes three different roles involved in data protection (see Art. 4GDPR):

- The *data subject* is the person whose data are processed and who needs adequate protection. This may include (individual) customers, employees, visitors to a company website, and many other individuals. In the current context, the data subjects will usually reside in Europe and have no direct relationship to the companies under consideration.
- The *controller* is the entity that is responsible for how data are processed within an organisation and may be the customer of the software development organisation discussed here. To a limited extent, software development organisations may also be the controller, for example regarding personal data collected via its website aimed at European customers or when offering Software as a Service (SaaS).
- The *processor*, finally, is the entity that does the actual processing of data, following the rules set by the controller. The processor and controller may be the same entity, in which case one usually just talks about the controller, or they may be different entities, in which case a contract between the two is required to ensure that the controller does actually control the

processing, and the processor follows the rules set by the controller.

In the current context, the software development organisation will in general be a supplier to the controller or the processor, and therefore only indirectly be affected by the GDPR. If, however, the software organisation does not only develop the software but also run a service based on this (or other) software, then it may itself become the controller or processor, which is, however, outside the scope of this paper.

GDPR requirements for non-EU software development organisations

This section summarises the main requirements set by the GDPR and their relevance on software development organisations working for EU customers.

Data protection regarding cooperation with European customers

When working for a European customer, an organisation will usually collect a certain amount of basic personal data about its contacts at the customer organisation, at a minimum names and contact details such as e-mail addresses and telephone numbers. While these data will in general not be critical, the individuals concerned must at least be informed about the fact that their data are stored and about their legal rights (Art. 12, 13 and 14 GDPR).

A second set of requirements concerns the company's website. If the development organisation's website addresses (among others) EU customers, visitors to the website must be informed about the personal data collected about them via website logging, the use of analysis tools such as Google Analytics, cookies or similar, the processing performed on these data, and the data subject's legal rights.

Data protection in software development

The requirements on data protection as defined in the GDPR (or other, similar regulations) lead to a number of requirements on software developed for use by companies controlling data of individuals in the EU. The current section will discuss these requirements in some details. However, a major challenge in this context is that customers will not always state these requirements on their own, and sometimes are not even keen on satisfying them in the first place. In such cases, it is important for the software development organisation to at least know and point out these GDPR requirements so that one can discuss about how to implement them.

Addressing the GDPR principles

The core ideas of the GDPR are presented in Article 5, which defines a set of principles that need to be satisfied in the processing of personal data:

Lawfulness, fairness and transparency: In particular, these principles state that the processing of personal data is forbidden unless one of six lawfulness conditions defined in Art. 6 GDPR is satisfied, such as consent by the data subject or "processing is necessary for the purposes of the legitimate interests pursued by the controller". This principle has to be addressed by appropriate organisational steps, while the software and its developers have very limited influence on their satisfaction.

Purpose limitation: It requires that personal data may only be used for the purpose for which they were originally (and lawfully) collected. In this case, the software and its developers have some but limited influence. The main requirement here states that before using personal data collected in one data base by some other application or for some other purpose, software development needs to ensure that the customer of the software development organisation (the controller or processor) is aware of this new use and confirms that this is legitimate.

Data minimisation: The use of personal data is to be reduced to the minimal extent necessary for its purpose. Storing personal data "just in case ..." is forbidden, which can become a serious problem in the context of big data. To implement this principle, the software development organisation needs to limit the personal data collected, stored and processed to the minimum genuinely needed for the purpose of the software. For example, if the data are only used for calculating aggregated parameters such as average or maximum, there is no need to include any identifying information and the data should be stored in an anonymised format. Any attributes that are not needed or used must not be collected and stored, even if they might possibly be useful for some future data analysis.

Accuracy: The controller and / or processor need to ensure that the personal data are accurate and up-to-date. Accuracy of personal data can be achieved or at least improved by standard data quality measures such as selection lists rather than free text fields for data entry where adequate, consistency checking of input data, and by providing functionality for correcting data once they are no longer accurate.

Storage limitation: It is closely related to data minimisation and requires that personal data are stored no longer than necessary. This principle is more complex to implement. It requires functionality to identify personal data that are no longer needed and to delete it or, if the data need to be stored to satisfy relevant legal requirements, to restrict access to these data so they can no longer be accessed for any other purpose.

Integrity and confidentiality: It requires that personal data are protected adequately to ensure that only people that are entitled to do so can read or write these data. To implement these principles, the software system

developed needs to incorporate adequate IT security measures such as user authentication and user permissions, typically based on defined roles.

Accountability: Finally, requires that the controller does not just conform to the above principles but is able to demonstrate this compliance. The task of the software development organisation mainly involves providing appropriate documentation to support the controller in this demonstration.

Addressing the rights of the data subjects

In addition to the principles described above, the GDPR defines a number of rights of the data subject. These rights lead to additional requirements on software systems, since many of them can only be implemented by incorporating suitable functionality into the system. To some extent, these rights look – though from a different angle – at the same topics as the principles above, thus leading to the same or at least similar software requirements. The main important rights in this context are:

Right of transparent information (Art. 12 GDPR): Data subjects are entitled to information about the handling of their personal data in a transparent and easy-to-understand way. This right does not imply any requirements on the software itself but on its documentation, which must allow to easily identify the personal data that within the software system.

Right of information (Art. 13, 14 GDPR): Data subjects have the right to be informed about the processing of their personal data. Like the previous right, this leads to requirements on the documentation rather than the software itself.

Right of access (Art. 15 GDPR): Data subjects have the right to ask for the personal data stored about them, and the processing performed on these data. This right implies that the software system must provide the functionality to identify and report all stored data about any specific individual. While this may be reasonably easy for structured data, for example in a typical customer data base, it can become extremely complex when unstructured data are involved, such as text documents, emails etc. Nevertheless, without such functionality a controller or processor organisation will not be able to answer the relevant questions by data subjects as required.

Right to rectification (Art. 16 GDPR): Data subjects have the right to request the correction of incorrect data stored about them. Implementing this right builds on the functionality needed for the previous rights. Again, this in general is not too difficult for structured data but can be very difficult for unstructured data. Additionally, there is the challenge to ensure that every instance of redundant data is corrected, possibly across multiple systems.

Right to erasure (sometimes also called the "right to be forgotten"; Art. 17 GDPR) and right to restriction of processing *(Art. 18 GDPR)*: Data subjects have, under certain conditions, the right to request that personal data about them is deleted or if for some reason this is not possible, for example because the data need to be stored for legal reasons to restrict the processing of these data. Probably the best known example of what the right to erasure involves is the Google search engine, where this right was first introduced by the European Court. Under certain conditions, Google must ensure that entries found by the search engine and concerning a certain person are not shown to users if this person requests so. In the case in point, the individual concerned had gone bankrupt many years earlier, and requested that information about this bankruptcy no longer be shown in Google searches. Expressing the right to erasure as general software requirements, there must be a function that allows the deletion of personal data and going beyond simple deletion, there must be a "black list" of data subjects for which (selected) personal data will not be processed in the future even if that would usually be the case. For example, it must be ensured that a person that has asked to be removed from the list of potential customers receiving marketing material will not be added again later based on a different source. Similarly, if data may not be deleted, at least access to these data must be restricted.

Right to data portability (Art. 20 GDPR): Data subjects have the right to transfer data provided by them to a different processor, e.g. if they want to move to a different provider for a certain service. The right to data portability is another challenge to software development, and the exact requirements resulting from it are still under discussion. To implement it, the software needs to provide some export functionality for the personal data that were provided by a data subject, in some structured, common and electronic format. Unfortunately, in many cases it is unclear what counts as an adequate format for the data export. Similarly, the systems need to have suitably import function as well, which may be even more difficult if there is no standard format to import.

Right to object (Art. 21 GDPR): Data subjects have, under certain conditions, the right to object to the processing of personal data about them. This leads to essentially the same functional requirements as the rights to erasure and to the restriction of processing, even though there are slightly different legal triggers when this functionality is needed.

Addressing other GDPR concepts

In addition to the principles and rights listed above, there are many more requirements in GDPR, some of which are relevant in the current context and will be discussed in the following.

The most important requirement on software development asks for data protection by design (Art. 25 (1) GDPR), which does not state any requirements on the resulting product but on the development process. Data protection by design requires that the principles and rights described above are addressed from the start of designing a process and the software to support it, rather than trying to add data protection later, which will usually be far less effective and efficient. Put simply, data protection by design states that data protection needs to be addressed across the entire development life cycle, in particular in the early stages of requirements analysis and design. An important step is to analyse from the start whether any personal data are genuinely needed for the purpose under discussion, or whether it is sufficient to work with anonymous data or at least with pseudonyms. If possible, pseudonyms or better anonymous data must be used. Using the similar term privacy by design, Cavoukian (2011) described an approach to implement data protection by design.

A related requirement is data protection by default (Art. 25(2) GDPR), stating that (software) systems must be configured such that privacy is the default and the user may change these settings to explicitly allow less privacy, rather than vice versa. The rules on automated decision-making and profiling set a limit to the usage of software systems (Art. 22 GDPR). Although decision-making solely based on automated processing is allowed in general, any data subject is entitled to obtain human intervention in such decisions, for example decisions about whether or not an individual is considered credit-worthy. Regarding software development, this implies that whenever any automated decision-making is performed, there must additionally be a possibility to perform this decision manually: the organisation must set up suitable processes, and the software must allow to manually over ride the standard automated decision making.

CONCLUSIONS

As the current paper shows, software developers need to take the GDPR into account even if the software is developed outside the EU. The GDPR requirements can be split into three groups. First and most important, there are a number of requirements on the software products as listed above. Second, the GDPR states requirements on the software development process, described as data protection by design. The third group consists of requirements on the cooperation of the software development organisation with its customers, in particular the adequate handling of customer data. Additionally, there will usually be further requirements derived from the local data protection legislation the details of which depend on the location where the software development organisation is based. When implementing these requirements, it is important to remember that data protection is concerned with the protection of legitimate interests of individuals, called data subjects in this context, and the GDPR as well as other data protection legislation above all document how to safeguard these legitimate interests, even though this will sometimes make live more difficult for the processor and the software developers supporting him.

REFERENCES

APEC. 2017. *APEC privacy framework (2015)*. Publication Number APEC#217-CT-01.9. Available at https://www.apec.org/Publications/2017/08/APEC-Privacy-Framework-(2015)

Cavoukian, A. 2011. Privacy by design. In *The 7 foundational principles*. Information and Privacy Commissioner of Ontario. Available at https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf

GDPR. 2016. 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). Available at https://eur-lex.europa.eu/eli/reg/2016/679/oj

Greenleaf, G. 2013. Nepal's unknown data privacy law: No Shangri-La, but a first for South Asia. International Data Privacy Law. **3**(4). Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2326799.

Magee, T. 2018. China's data privacy standard came into effect this May- inspired by GDPR. Available at https://www.computerworlduk.com/data/how-chinas-data-privacy-law-was-inspired-by-gdpr-3678918/