A Practical Look into GDPR for IT

Part 3

Abstract

This is the third and last article in a short series about the new EU General Data Protection Regulation (GDPR). In this article we will discuss some consequences for IT systems deriving from the GDPR's requirements concerning the data breach management and the fulfilment of the citizens' rights on the management of their personal data.

In this last article on the GDPR and its impacts on IT, we will discuss two main subjects: data breaches and IT services to satisfy the citizens' rights.

Data Breach

The approach to IT Security of the GDPR is quite practical and assumes that IT security incidents and in particular data breaches concerning personal data relevant to the Regulation (Articles 33 and 34), are possible and will happen. A company is thus expected to:

- implement security measures to reduce the risk of a data breach;
- notify the supervisor authority in case of a data breach within 72 hours after having become aware of it, and describe to the supervisor authority the security measures taken to mitigate the effects of the breach;
- manage the security incident informing, in case of high risk to the personal data, the interested parties of the possible consequences of the breach and of the actions taken to mitigate it.

Notice for example that a breach can be effectively mitigated if the data is encrypted and the decryption keys have not been breached.

At first sight the GDPR requirements on data breach look reasonably simple, but at a closer inspection it turns out that they are not easy at all to implement. From an IT security point of view, it is possible to break them down in three major activities:

1. implement proactive IT security measures
2. monitor the IT systems for security breaches
3. manage IT security incidents.

We will briefly describe how these three points are related, in particular how the third one depends on the first two.

The first consideration is obvious, one needs to have informations about incidents to be able to manage them. Unfortunately it is true that too often an incident is reported from outside a company's IT, even from outside the company.[1] Still it is essential to implement a full monitoring of the IT systems so to be able to identify:

- violations of security policies
- attempts to breach the IT systems
- breaches of the IT systems.

Monitoring is not easy, it requires to collect a large amount of data and to analyse it, to correlate events and to evaluate their significance based on the peculiarities of each IT system. Indeed many events which could be normal for one IT system, could instead be indicators of compromise for another IT system. The complexity of this task has led to the development of a full industry of SIEM (Security Information Event Management), Security Analytics, SOC (Security Operation Centres), and recently also the use of machine learning, artificial intelligence and cognitive security.[2] Still implementing those three actions above is not trivial, requires to collect extremely large amount of data and can be rather expensive.

Moreover, managing security incidents often requires much more from the monitoring system than receiving the alert of a breach. Indeed a breach investigation requires to be able to reconstruct all events happened in the past, in some cases days, weeks or months in the past, that have led to the breach, and all the activities of the perpetrator of the breach or activities which could have helped the perpetrator even involuntary.

Given a breach event, one should reconstruct what has happened before it and after it. One should also verify if there have been previous or successive attempts, failed or successful. All this requires

---

1 See for example the 2017 Verizon's DBIR report [2] which reports that 27% of breaches were discovered by third parties.

2 For a well known public example see IBM's Watson for Cyber Security [3].

to collect and maintain for sufficient long time enough data to be able to perform these analysis and investigations. The main problem is that a priori it is not known which data will be needed for these investigations and for how long it should be kept, so in principle one should collect and keep sufficient monitoring data to be able to reconstruct any event on any IT system for a long period of time, which is obviously impossible technically and economically.

But, or according to the GDPR, when a data breach happens, a company must be able to manage it and effectively mitigate the damages. First of all this requires that the company has in place an incident management process which defines who should do what, which are the responsibilities, actions to be performed and reported etc.

Besides organizational measures and personal, technical skills, to effectively manage a data breach it is also very important that some proactive security measures are already in place. The simplest way to explain this point is possibly by an ideal (not realistic) example: assume that a company has a completely flat internal IT network, each device connected to the internal network can communicate without any filter to any other device in the internal network. Assume also that there has been a breach on one device and one of the tools used to perform the breach by the attacker is a worm which tries to propagate by itself to machines directly connected. How can the breach be effectively mitigated in this situation from a network point of view? Very hard to say.

Assume instead that the network had been previously partitioned in security zones with firewalls and other security devices to filter the traffic between the security zones. In this case it is possible in very short time, that is "effectively", to isolate all security zones so to prevent the further distribution of the worm between them. Then one could proceed to verify which zones are infected and which are not, and to clean up the infected zones.

This fictitious example shows us that to be able to effectively manage data breaches and mitigate their damages, it is necessary to introduce proactively IT security measures which can be fundamental tools during the incident management. Indeed one of the most frequent points in the "lessons learned" of an IT security incident post-mortem evaluation, is the lack of the appropriate tools, configurations or IT design/processes which made the incident management not effective or even almost impossible.

Citizens' rights

The GDPR has also requirements related to Citizen's rights in Articles 15 to 22 that IT systems need to fulfil. Here we will not discuss all Articles on citizens' rights of the GDPR, but only those which seem more relevant to our discussion.

First of all, it should be noted that the GDPR requires that the citizens' requests which correspond to GDPR's rights, must be satisfied "*without undue delay and at the latest within one month*" (Recital

59). This implies that IT processes and functionalities must be already in place to satisfy these requests, which otherwise would be quite difficult to satisfy only with manual activities.

We will now discuss some of the rights individually, starting from the easiest ones.

Right to rectification (Art. 16)

This is possibly the simplest right from an IT point of view, if some data has been inserted in some IT systems, in the same way it should be possible to modify it. But there are a few corner cases which should be considered and which require that the IT systems' behaviour is aligned with the business and person's expectations. For example consider the cases of change of address, marital status, sex or even name: in which cases a 'new person' should be created in the IT systems independent from the existing one? What should be traced of the modifications? Should the old address be retained together with the new one or only the new one should be kept? In the cases of change of name or sex it is possible that the business would like to keep the same 'person' within the IT systems tracing the modifications, but that the individual would like instead to start a new 'history' of her/his personal information. In the worst cases, this could lead to the development of quite complex IT functionalities.

Right to erasure ('right to be forgotten', Art. 17)

We are not concerned here with the conditions or the limitations on the right of a person to request to erase her or his personal data from the IT systems of a company; there can be legal and/or contractual clauses which can void or postpone this right, as for example the need to wait for the expiration of a contract. We assume that it has already been stated that the GDPR data of a person has to be erased from the IT systems. Some of the issues that can arise are:

- the need to know where the GDPR data of the person is stored, in which applications and IT systems
- the need for all of these applications and IT systems to have functionalities which allow to erase the GDPR data of a person without losing consistency and integrity; for example in an online shop the erasure of a customer and of all her or his transaction history can invalidate the consistency and integrity of the products' availability and accounting
- the need to erase the personal data also from backups and offline storage
- the need to request the erasure of the data also from IT systems of third parties which could have manage it on the behalf of the company.

In particular the second and third point above could be difficult to implement. The second point could require some custom development for all applications managing GDPR data. Moreover current technologies for backing up data are not optimized for deleting only some records from the backups, indeed many traditional backup applications only allow to erase a full backup or nothing.

A possible approach to this problem, to which we will come back later on, is to introduce a new, centralized application devoted only to manage GDPR data. All other applications do not store GDPR data but acquire it from this centralized application which can then manage also the erasure in a single place. Moreover this application can have special backup processes which are organised by person so that it would be possible to erase also from the backups the personal data of a single person.

Right of access by the data subject (Art. 15) and Right to data portability (Art. 20)

We consider both Articles 15 and 20 together since from the IT point of view it should be easier to design and implement a solution for both at the same time. The first thing that these Articles require is that a person is allowed to know if her or his data is managed by the company and who can access it, but also how the data is managed, for which purposes, for how long is stored (retention period), and finally the person has the right to have one copy of all her or his data.

There are two aspects for what concerns the access to the data:

- the roles of the company personnel (including the IT system administrators) who can access the data, for which purposes and in which way
- who has actually accessed the data, when and how she or he has treated it.

The first point requires the company to describe in policies and procedures who can access the GDPR data and how. This then is summarised in the information for the persons. The second aspect requires instead to implement monitoring of the IT systems to log the access to the GDPR data by the company personnel, and depending on the sensitivity of the GDPR data this can be done at a summary level (tracing only login and logout on the IT systems) up to the finest detail (tracing each single access to the GDPR data both for reading and writing).

The right to data access and data portability can be in practice quite complex to implement. The GDPR requires (Article 20):

1. *The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without*

hindrance from the controller to which the personal data have been provided, where:

a. the processing is based on consent [...]; and

b. the processing is carried out by automated means.

2. *In exercising his or her right to data portability pursuant to paragraph 1, the data subject shall have the right to have the personal data transmitted directly from one controller to another, where technically feasible*.

The data portability requirement practically includes all possible data that a person has provided to an IT service. In [4] the WP29 provides some examples like:

- a person's current playlist from a music streaming service to find out how many times he listened to specific tracks in order to check which music he wants to purchase on another platform
- a person's contact list from his webmail application to build a wedding list
- a person's emails from his webmail application to send them to a secured storage platform
- a person's details of his bank transactions to send them to a service that assists in managing his or her budget
- the titles of books purchased by an individual from an online bookstore.

Obviously, each company will have to evaluate and identify which are the personal data subject to the data portability requirement, the Guidelines [4] of the WP29 give some indications on how to do this. The job of the IT department will then be to implement functionalities to make this possible. But this can be quite complex.

Interpreting a little what could be the idea behind the GDPR, a person can expect to find a button on a company's website which allows to download all her or his personal data "*in a structured, commonly used and machine-readable format".* To make this possible, we should expect that there will appear international standards for personal data formats based for example on XML or XHTML, which will allow both the direct access of the individuals to the data and the exchange with other companies.

On the company's website there should also appear forms which allow the person to transmit her or his personal data to another company. Obviously in this case the two companies must agree in advance and have established a technical communication channel to exchange this information and the associated legal requirements implied by the transfers. For this to happen many issues have to be cleared not only at the technical level, eg. which technical standards to adopt, who is responsible for what in the communication and in the security of the exchange (mostly anyway in charge of the sender), but also at the commercial, business and legal level.

On the other side, to be able to provide these services, there should be IT services which:

- are aware of all GDPR data
- are able to collect it
- are able to assemble it in a common format
- and are able to export it.

One possible technical approach to this is to implement an IT service which is aware of all other IT services managing GDPR data and that it is able to access all this services and retrieve the data. In practice this is often extremely complex to implement and to maintain up to date since new services and new GDPR data can be added at any moment.

An opposite approach, already mentioned in a previous section, is to design a new application dedicated to manage only GDPR data. All other applications will get GDPR from this central point either in real-time or by periodic updates. This approach can have many points in favour like the following:

- all GDPR data is stored in a common format
- it is easy to export an individual personal data and satisfy both the requirements of data access and data portability
- it is easy to maintain all GDPR data current and up to date
- it is possible to design specific backup policies which in case will make it possible to satisfy the right to erasure
- it can help (in particular in case of real-time access) on tracing accesses to the GDPR data.

On the other side, it introduces another application in the IT service landscape and another possible point of failure for data availability, confidentiality and integrity.

Right to object (Art. 21) and Automated individual decision-making, including profiling (Art. 22)

Finally we consider the Articles 21 and 22 which introduce the right for an individual to object to have her or his personal data managed by some processes and, in particular, to automatic, that is IT, processing. Two particular significant statements in these Articles are the following:

- *Where personal data are processed for direct marketing purposes, the data subject shall have the right to object at any time to processing of personal data concerning him or her for*

such marketing, which includes profiling to the extent that it is related to such direct marketing

- *The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.*

Since as of today marketing campaigns, big data analytics, data warehouse, business intelligence and customer relationships are all managed by IT services, all these services will have to provide functionalities to exclude from data processing for direct marketing and/or automated processing including profiling, all individuals who request it.

Again these could be quite relevant requirements at the technical IT level since some applications could need major modifications or introduction of new features to be able to satisfy them. But also at the business level there should be modified or new processes to manage these situations and these particular rights of the individuals.

This concludes our practical look into GDPR for IT: this new legislation will induce in any case many modifications and changes in the current IT systems which hopefully will not turn out to be only legal burdens for the companies (and ultimately for their customers who in some way will have to pay for them) but also will create better, more secure, easier to use and more trusted IT services for everybody.

References

[1] European Union, “General Data Protection Regulation 2016/679”, http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG&toc=OJ:L:2016:119:TOC

[2] Verizon, “2017 Data breach Investigations Report - 10th Edition”, http://www.verizonenterprise.com/verizon-insights-lab/dbir/2017/

[3] IBM, “Watson for Cyber Security”, https://www.ibm.com/security/cognitive/

[4] Article 29 Data Protection Working Group, “Guidelines on the right to data portability”, http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=50083

Andrea Pasquinucci (PhD CISA CISSP)