Big Data 101

Felicitas Ronneberger, Tugrul Aras, Silas Weber and Philip Hofmann

Big Data: Technology and Law
University of Zuerich

Abstract

Big Data 101 is a native android app, specifically tailored towards insurance companies residing in Switzerland. The app touches on the subjects of the technological and legal aspects of data such an insurance company might hold, explicitly the topics of big data and the general data protection regulation (GDPR), by explaining the matter in static content form. It strives to help the user getting a theoretical understanding of these issues in regards to GDPR compliance. To be informed about applicable legal and technological developments like court rulings, the app includes a news section where articles are fetched from our web server. These articles are submitted by the development team or trusted third parties using the web interface.

Introduction

The countdown to compliance has begun. The application of the GDPR in all EU member states and Switzerland will start on 25 May 2018. With the help of our app 'BigData101' the insurance industry in Switzerland will be provided with an overview of the impact the GDPR will have with regards to people in manager positions within that industry. The insurance industry will need a greater command over the data it is holding, why it is being held and how long it is being held for. This will require a seismic change of attitude for many companies. Fines, which can now be as much as 4% of annual worldwide turnover, will mean that data protection will need to be on the boardroom agenda.

Native Android App

The android app consists of static and dynamic content. The static content involves a welcome screen at the beginning where we will mention the name of the app and names of the creators but more importantly the legal disclaimer is displayed. This is the screen the user is brought to automatically at the start of the app.
The thematic content is divided by chapters. It guides the user towards an understanding of the technological as well as the legal aspects in regard to big data and the GDPR. This content is specifically tailored towards insurance companies in Switzerland. Additionally, we recommend a course of action for insurance companies that wish to comply with the GDPR.
The dynamic content is the news update part that will get fetched from our server and displayed to the user. This is to keep the user updated about development in both legal aspects, which might involve new court rulings, and technological aspects where and if applicable.
A technical difficulty was dealing with fetching the articles from the server while keeping the user interface free to navigate. To solve this we implemented an asynchronous request with callbacks to update the user interface. Additionally, to enable the app to deal with a growing amount of data, we chose to implement the more sophisticated Recyclerview instead of a normal ListView.

Legal Aspects

Introduction

The consequences for a breach concerning data protection laws are substantial. Therefore, distinctive GDPR trainings to sensibilise management and staff are indispensable. Part of this interdisciplinary project is the app Bigdata101. In a first step, the legal issues concerning the app itself will be explained, furthermore, legal aspects about the app's content and at last the relevant changes for insurance industry, including the rights and duties according to the GDPR. The application setup lets employees detect whenever the GDPR is of relevance and what consequences might occur in case of a violation. Its focus is on the GDPR. Therefore, it is considered that the companies operate in adherence to local and federal laws and regulations. The comparison between the GDPR and the Swiss provisions (also considering the draft of the Swiss Data Protection Law (DPL)) shows that they have several points in common. Some points will be mentioned as part of the GDPR even if they are already part of the federal law. Some legal issues will be repeated because of thematic overlapping and inner cohesion within the GDPR. Theoretical sessions about the main elements of the GDPR will be explained on examples or procedures/instructions. Legal knowledge and news are displayed in the news

section. The app does not replace internal classroom trainings, moreover should it provide employees a first glimpse of the GDPR topic. Legal obligation will be waived with a disclaimer which the user has to accept when accessing the tool. The content in the app is for information purposes only without any legal correctness.

Switzerland

The GDPR will not only apply to insurance companies and their data controllers and data processors established in the EU but also to those who are concerned with EU resident's data or any data exchanges inside the EU [5]. Despite the fact that Switzerland is not part of the EU, the GDPR, which goes into effect on 25 May 2018, will have a direct impact on the insurance sector. In light of the horrendous fines it is even more important to provide information about fundamental differences and characteristics of the GDPR including risks and benefits concerning Swiss insurance companies.

Insurance industry

Brokers and insurers possess massive databases full of customer data. In the event of a liability case, client data is received and processed. With that number of businesses and services operating across borders, insurance companies, respectively managers of those companies, have to comply carefully with the GDPR not only because of their liability as part of the administrative board Art 716a and 754 Swiss Obligation Law. Risk parameters are being valued according to certain client data like personal health data. Additionally, interdisciplinary collaboration in fields such as diagnosis and medical treatments between doctors, hospitals and other insurance companies is key. Occasionally, insurance companies are being asked to reveal client data. Employers for example are seeking information regarding the health status of their employees to get information about when their employees will be able to return to work after an incident. Some patients might also have their domicile in the EU and their data is requested - this means that many obstacles exist on a daily basis.

The application setup lets employees detect whenever the GDPR is of any relevance and what consequences might occur in case of a violation. Regarding the app itself there will be no data exchange. Customers will only be able to navigate, not to edit. To see the content of the app, a disclaimer will be written regarding the principles "privacy by design and default" and which must actively be clicked and confirmed.

Theoretical sessions about the main elements of the GDPR will be explained with the help of examples, instruction and cases.

Law aspects concerning the app

There is no further exchange of data (such as customer data) required for technical operations. The news section will be protected by a password, which is only given to trusted third parties such as specialists or lawyers. Concerning the content, a legal disclaimer which implies the exclusion of liability shall apply for all types of damages and in particular for damages that could result from incorrect contents, loss or deletion of content.The disclaimer must actively be selected to continue. Including the legal disclaimer, within the terms and conditions rightfully accepted by the user, is a viable option. In such cases the fundamentals regarding the general terms and conditions would be applicable. The most reliable way to secure that a disclaimer is being legally accepted by the user is to obtain their explicit consent.[2] Eventually, many websites use a designated pop-up field that users need to accept before browsing the website. As already mentioned it will be part of the app to accept the legal disclaimer before being able to continue.[10]

The apps's content

Terms and definitions In order to better understand the scope and key issues of GDPR the app user should first get familiar with some frequently used terms. So there will be an introduction about a key selection of definitions based on Art. 4 GDPR. The other definitions can be viewed under Chapter 1 - Art.4 GDPR. Taking into account that the provisions relating to GDPR will be newly introduced into European law, it is not predictable how the courts will decide on actions, interpretation and legal disputes. Cases can only be added at a later date.

- personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

- processing' means any operation or set of operations which are performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;

- Controller: natural or legal person, public authority or agency (etc.) which determines the purposes and means of the processing of personal data Processor: natural or legal person, public authority or agency (etc.) which processes personal data on behalf of the controller

- filing system' means any structured set of personal data which is accessible according to specific criteria, whether centralised, decentralised or dispersed on a functional or geographical basis;
- Recipient: natural or legal person, public authority or agency (etc.) to whom the data is disclosed
- Third party: natural or legal person, public authority or agency (etc.) other than the data subject, controller or processor who is authorised to process the data under direct authority of controller or processor

Further legislation:

- KVG Federal Health Insurance Act of 18th march 1994, LS 832.10
- DPL Swiss Federal Data Protection Law of 19th June 1992, LS 235.1

Scope Art. 1 GDPR states the protection of fundamental rights and freedoms of natural persons, in particular their right to privacy and expedites free movement data within EU. [5]The regulation is applicable to the processing, completely or partly by automated means such as digital databases. In addition, the processing of personal data by any other means is also regulated by the GDPR when the data is included in a filing system Article 2(1) GDPR. Importantly, Article 3(2) GDPR states that the regulation is applicable when controllers and processors are not established in the European Union but process personal data of individuals who are in the Union. Such processing activities must relate to the offering of goods or services, for a payment or for free, to these individuals or to the monitoring of the behaviour on European level (also webtracking), Article 3(2)(a) and (b) GDPR. It is questionable whether the GDPR is applicable if it is not foreseen to participate in the EU-market.[5] [1] [4]

In association with insurance data, various scenarios for an application of the GDPR are possible, for example if the company offers international insurance policy services or personally mentors customers who are settled in the EU.[3] In the event of a liability case, client data is received and processed. Risk parameters are being valued (maybe as part of monitoring) according to client and damage data. Additionally, interdisciplinary collaboration in fields such as diagnosis and medical treatments between doctors, hospitals and other insurance companies is key. Occasionally, insurance companies are being asked to reveal client data. Employers are seeking information regarding the health status of their employees and when they will be able to return to work after an incident. Patients might have their domicile in the EU and their data is requested - many obstacles exist on a daily basis.

There is an excellent mnemonic aid: The GDPR applies if the processor or controller is settled in Switzerland and the antagonist is settled in the EU (for example in case of outsourcing) Does your company...

- have customers with their seat in an EU country and do you offer them services (like international insurance policy)
- have an establishment within the EU
- monitor the behaviour of EU citizens within the EU (this can also be possible by data processing on an app or website)
- process personal data of employers established in the EU

If one of the mentioned conditions is satisfied, the GDPR applies.[5][1]

stay abroad The case of insurance protection during a time-limited stay in a foreign country and paid services of health insurance

cross-border commuter Persons who live in any EU member country but work in Switzerland must insure themselves and their non-active family members in Switzerland. (Federal Health Insurance Act/KVG)

seasonal workers Persons returning back to the EU and having some claims on insurance benefits (for example health insurance),

Keypoints Under the GDPR, the data protection principles set out the main responsibilities for organisations which are quite similar to Swiss law and are stated in Art. 5 paragraph 1 lit. a-f & 2 GDPR : [5]

- lawfulness, fairness and transparency'
- purpose limitation
- data minimisation
- accuracy
- storage limitation
- integrity and confidentiality
- 'accountability

Accountability GDPR brought some big changes such as in accountability. In Article 5 (2) it is stated that your organisation is responsible for and shall be able to demonstrate compliance with the other principles from paragraph 1. To meet the requirements of accountability, it is important for you to install appropriate organisational and technical measures. In some cases there are measures which you must or can take, like:[1][3]

- deciding on a data protection officer;
- keeping documentation of your processing activities;
- adopting and installing data protection policies;
- taking a 'data protection by design and default' approach;
- preparing written contracts with organisations that process personal data by your order;
- recording and, where necessary, reporting personal data breaches;
- installing proper security measures;
- registering to certification schemes and considering relevant codes of conduct; and executing data protection impact assessments for uses of personal data that are likely to be categorised as high risk to individuals' interests.

Records of processing activities Art.30 GDPR There are precise provisions about documenting your processing activities in the GDPR.[1][3] You have to consider that some actions such as data sharing, retention and processing purposes have to be kept. With the help documentation, you can improve both your data governance and comply with other aspects of the GDPR. Documentation obligations must be met by controllers and processors. For small and medium-sized organisations, documentation requirements are limited to certain types of processing activities. Your records have to reflect the current processing actions and they also have to be kept up to date. Example:Data collection regarding CRM (Customer-Relationship-Management)

Right to erasure Customers have the right to have personal data erased, which is given to them by Art.17 GDPR. Under Swiss law, this specific right to erasure is also known as 'the right to be forgotten'.[7] It is also not explicitly regulated in Swiss law but can be provided in accordance to the principle of Art. 4 DPL. If customers want to start the process of erasure, they have to make a request verbally or in writing. You have one month to respond to a request. The request must then be examined. The right to erasure only applies in certain circumstances and is not absolute. There are also other ways in which the GDPR asks you to consider whether to delete personal data or not. The company must comply without "undue delay" and at no cost to the requesting party.[3]

Case An insured person sends a request to erasure personal data.

- Step 1 Check if the data needs to be erased
- Step 2 Find out where the requested data resides.
- Step 3 Evaluate all source data, reprocessed and extracted data systems.
- Step 4 Evaluate data which was given to third party systems (outsourcing)

Right to data portability A new right is introduced with the Art.20 GDPR where customers are allowed to obtain and reuse all kinds of personal data for their own purposes across different services.[8] Without obstacles in usability, this article also allows you to move, copy or transfer personal data easily from one IT environment to another in a secure manner (to keep data on their personal devices). Data portability helps taking advantage of applications and services which can use this data to optimise consumption behaviour like spending habits or finding more attractive deals. This has a major impact on technical aspects which data controllers will need to ensure that their systems, connected products, applications and devices, which collect and store information on data subjects, also have the added functionality of porting and transmitting data.[1] In some cases, this will require controllers to tweak or redesign some systems, products, applications and devices. Furthermore, the new porting functionality must export data in a structured, commonly used and machine-readable format so that reuse of the data is possible. The GDPR does not establish a general right to data portability for cases where the processing of personal data is not based on consent or contract. Cases: A person who is voluntary insured (right to portability is given), A person who is under the compulsory insurance (for example as part of the health insurance, Federal Health Insurance Act/KVG) has no right to portability because the processing of personal data is not based on consent or contract but based on law.

Privacy by design and by default Under the GDPR, you have a general obligation to implement technical and organisational measures to show that you have considered and integrated data protection into your processing activities.[7]Privacy by design aims to protect the users' privacy and give them control about their data by measures of the design stage.The privacy by default approach states that systems should be designed to always ask for permission to process/collect data.[3]

This means that data controllers will need to take account of the following:

- Step 1 Build the right to data portability and the right to be forgotten directly into the software.
- Step 2 Identify the risks concerning the rights and freedoms of individuals.
- Step 3 Identify the data flows of products and services
- Step 4 Develop new standard processes and templates which are compliant to the GDPR

Processor & Controller The application of the GDPR possibly establishes a fairer balance between data controllers and data processors considering the continuing role of data controllers.[3] On one hand data controllers had to deal with their own data protection compliance obligations and on the other hand with the actions of their processors, which was seen in an unfair relation. So now there is a possibility for brokers and insurers, with some exceptions, to enable data controllers themselves to minimise the liability.[3]

However, long-term negotiations about contract matters could occur since data processors will inevitably need precise contractual provisions regarding:

- the agreed relationship between the actors with respect to each aspect of the processing action;
- data controller and data processor subject to their responsibilities(Art.24 GDPR).
- the ensurance of their own compliance with GDPR obligations in connection with receiving specific processing instructions.

Such obligations could slow down data processing agreements and negotiations by making them too complex. So, you should be prepared for long-term negotiation processes. You are responsible for reviewing data classifications to ensure that data, which will now definitively be deemed personal data, is subject to the appropriate protections. •Ensure all data from which an individual can be identified (which will include location data and IP addresses) is covered by privacy policies and fair processing notices.[1]

- You need to review all data processing arrangements to ensure the contracts contain:
- all of the requirements set out in Article 28; and
- appropriate risk allocation of liability for data breaches between data processors and data controllers.

Furthermore, you should review contracts with a prioritised system taking into consideration sensitivity and volume of personal data that is processed.

Security of processing 32 The 'security principle' as a key principle of the GDPR sets the scale for you in processing personal data securely in a manner of 'appropriate technical and organisational measures'.[9] To stick to this principle, you will need to take into account topics like organisational policies, physical and technical measures and risk analysis.

- Furthermore, you have to consider extra conditions about the security of your processing. This applies to data processors as well.
- You should consider varying your measures. Where conventional, you can for example use pseudonymisation and encryption.
- Depending on your circumstances and your processing risks, you can take the state of the art and costs of implementation into account when choosing the right measures.
- In the case of a technical or physical trends, your chosen measures must be able to grant you restoring access and availability to personal data in due time.
- The 'confidentiality, integrity and availability' including the process of personal data in your systems and services have to be complied with the measures you choose.
- Besides, in terms of testing the effectiveness of your measures, you have to consider that you have installed convenient processes. Hence, you should conduct any kind of technical and organisational optimising.[3][1]

Reporting Art 33 GDPR With the implementation of the GDPR there will be a further obligation on all organisations, which is reporting specific kinds of personal data breach to the appropriate supervisory authority. Where manageable, you have to report the breach within 72 hours of noticing it.[3][1]

- You also have to brief those individuals immediately on the breach, whose rights and freedoms could be in a high risk of being influenced.
- Independent of the fact whether you are obliged to inform or not, you also have to keep a record of any personal data breaches.
- It is important for you to have solid breach detection systems, investigation and internal reporting procedures. Therefore, decision-making processes can be fostered whether or not you have to inform appropriate supervisory authority and the affected individuals.[1]

Data protection impact assessment Art. 35 (DPIA) If your organisation wants to identify and minimise the data protection risks of a project, the data protection impact assessment (DPIA) is a big help.[9] Your DPIA has to:

- characterise the context, scope, purposes and nature of the processing;
- detect and deal with risks to individuals
- evaluate proportionality, compliance and necessity measures;and
- detect further measures to minimise those risks.[9]

Data protection officer Art. 37 GDPR If you are a public authority or if you execute particular types of processing activities, the GDPR sets out the obligation for you to announce a data protection officer (DPO).[6]

- There are specific tasks of a DPO. He/She takes the role of an advisor as he/she helps you checking internal compliance, updates you on your data protection duties, allocates you on the topic of Data Protection Impact Assessments (DPIAs) and is a general contact person for supervisory authority and data cases.[3]
- The DPO has to fulfill some features like showing the required knowledge about data subjects, properly resourced, independent and with competence to report to the highest management level.
- DPOs are a big support for you to ensure compliance and meet the obligations of accountability.
- In an existing organisation, the DPO can be appointed from a group of employees who are already working in that organisation or could also externally be selected.[1]

Consequences and sanctions

Fines The regulation lays out maximum penalties which differ depending on the type of offense. One of the significant impacts is that the GDPR increases the maximum fine amount – up to €20,000,000 or up to (what can also be higher) four percent of the company's annual "global turnover of the preceding year.[3] [1] The consequences for global giants (such as Amazon), which have total revenues, net of taxes in the billions are exorbitant sums. The amount of fine – if any – that is actually imposed will depend on a number of different factors. An administrative fine must be effective, proportionate and dissuasive. Article 83 of the GDPR addresses in detail the conditions for imposing administrative fines, and specifically names factors that are to be taken into consideration:[1]

- The nature, gravity, and duration of the violation
- The categories of personal data that are affected
- Previous violations
- Intent or negligence
- Actual harm done and efforts to mitigate the damage to data subjects
- Degree of responsibility of the controller or processor
- Certifications and adherence to codes of conduct
- Reporting of the violation
- Cooperation (or lack thereof) with authorities

In addition, the €20,000,000 maximum apply to the higher of two tiers of violations, which include more serious offenses, such as those pertaining to the rules for obtaining consent, data subjects' rights, rules governing data transfer, obligations to member states, and violation of an order. Namely for infringements of basic principles which are stated in Art. 5, 6, 7 and 9 GDPR.
The lower tier of violations has a maximum fine limit that is half that of the upper tier: €10,000,000 or two percent of annual turnover. Some violations that fall into this category include:

- Notification of a data breach to the data subject whose personal data was impacted
- Notification of a data breach to the supervisory authority
- Failure to properly designate a data protection officer (when required)
- Certain conditions surrounding obtaining a child's consent

Reprimands In addition to or instead of fines, Article 58 provides for the issuance of warnings and reprimands . You could also have your certification withdrawn, or be ordered to take action to carry out one or more of the obligations under the regulation[3] [1]

Technology

Besides the native android app, we implemented a REST server that talks to a database for storing and retrieving articles written by the users.
This section explains the technical architecture and the details of each component.

Tech stack:

- For the web page we used plain HTML with Bootstrap and JQuery.
- The server runs on the NodeJs run-time environment and was developed with the use of the Express web framework.
- For the database we use the NoSQL database MongoDB
- Our application is deployed on Heroku: `https://bigdata101.herokuapp.com/`
- The database is on mLab: (Login credentials needed to access the dashboard) `https://mlab.com/databases/bigdata101`

We chose those platforms because they offer free hosting of our web service and our database respectively. With Heroku it was possible for us to set up a deployment pipeline that

publishes a new version online as soon as a change in the master repository on GitHub is detected.

On the web page there is only an interface for sending a POST request to the server, but our API also supports (next to sending GET request obviously) sending UPDATE and DELETE requests. This means articles can also be modified or removed, there is just not yet a graphical user interface for those operations as we focused on the app development and not on a web client that fully consumes the API.

The endpoints for the operations mentioned above are the following:

- GET `/api/articles/:articleId`, fetches the article with the specified Id
- GET `/api/articles/`, fetches all articles
- GET `/api/articles/:category`, fetches all articles from the specified category i.e (Law, Technology)
- PUT `/api/articles/:articleId`, updates the article with the specified ID
- DELETE `/api/articles/:articleId`, deletes the article with the specified ID
- POST `/api/articles/` inserts a new article into the database

The POST and PUT requests expect a request body to be sent along. The request body must be in JSON format and should contain the following keys:

- articleObject, wraps the actual article
- category, should be either 'Law' or 'Technology'
- article, the text of the article
- author, who wrote the article
- pwd, the password for authentication

Example of a POST request body:

```
{
"articleObject": {
  "category": "Law",
  "article": "my article text",
  "title": "my title",
  "author": "my name"
},
"pwd": "123"
}
```

The technical difficulties during the server development phase were that we had to think about securing our API endpoints when we put the application on Heroku to restrict unauthorized access to the data in the database. To solve this, a password (123) is required to send a POST request. PUT and DELETE requests on the other hand are authorized by an authorization token in the request header.[10]

References

[1] https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr.

[2] Oliver Staffelbach et al. *Social Media und Recht für Unternehmen*. Zuerich, 2015, pp. 25–42.

[3] Rhiannon Webster et al. *The European General Data Protection Regulation, A guide for the insurance industry*. https://www.dacbeachcroft.com/es/gb/articles/2016/june/the-european-general-data-protection-regulation-insurance-industry/.

[4] Sergio Ceresola. "Aktuelles zum Datenschutzrecht: Neuerungen in der EU und in der Schweiz". In: (2018), pp. 177–178.

[5] Francois Charlet. *GDPR in Switzerland: 10 steps organisations should take*. https://francoischarlet.ch/2017/gdpr-in-switzerland-10-steps-to-take/. 2017.

[6] Yves Gogniat. *Brauche ich einen DPO in meinem Unternehem?* http://www.dieadvokatur.ch//publikationenFachartikel/Brauche-ich-einen-DPO-in-meinem-Unternehmen.pdf. 2016.

[7] Marc Langheinrich. "Mehr Datenschutz durch Technik?" In: *Digma – Zeitschrift für Datenrecht und Informationssicherheit* (2017), pp. 14–19.

[8] Michiel van Schaick. *GDPR Top Ten: Data Portability Legal obstacel or opportunity*. https://www2.deloitte.com/ch/en/pages/risk/articles/gdpr-data-portability.html. Zuerich, 2018.

[9] Ursula Sury. *Impact Assessment Informatik Spektrum*. http://www.dieadvokatur.ch//publikationenFachartikel/Privacy-Impact-Assessment.pdf. 2017.

[10] Rolf H. Weber. *E-Commerce und Recht Rechtliche Rahmenbedingungen elektronischer Geschäftsformen*. 2nd ed. 2010, pp. 14–19.