Lesson 8_Data_Protection

GDPR (General Data Protection Regulation)

GDPR is a very large regulation and it's the main instrument that we have today for the regulation of computing. It doesn't speak about AI because even if the regulation has been released recently it reflects a debate that has taken place in the previous years and is focused on the challenges emerging from the internet but there are many provision on AI that we can find in this text.

Personal Data

The key notion in the GDPR is the one of personal data. This notion is relevant not only for lawyer but also for computer scientist because when an item qualifies as personal data then you have all the rule that concern data protection that comes into play: you cannot process this piece of data unless there is a legal basis (consent by the individual, a contract), you have to inform the individual that his data has to be processed. So, estabilishing if a piece of data is personal or not makes a lot of difference not only in legal domain but also in computing.

> ‘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; that is when a piece of information which is accompained by other data referring to the same individual and in such a way the individual can be identified then the data is personal.

Two concept very important in this context are:

- **Anonymisation** is a process through which some pieces of data is removed from a dataset to make the individuals non identifiable
- **Pseudonymisation** is the process of keeping the pieces of data that identifies an individual into a separate file and substituting them in the original dataset with an identifier. In this case a person can be identified only by having access also to the separate file.

To what extent AI can change the notion of personal data?

The technological tools provided by AI may enable, to some extent, the repersonalization of anonymous data and maybe the **reidentification** of the individual to which the data is related. Then AI may **infer** further personal information from the personal data available.

Reidentification

AI, and more generally methods for computational statistics, increases the indentifiability of apparently anonymous data since they enable nonidentified data (including data having been anonymised or pseudonymised) to be connected to the individuals concerned. This because the reidentification is usually based on statistical correlations, which can be learned by ML models, between nonidentified data and personal data concerning the same individuals.

In fact, in 2016, journalists reidentified politicians in an anonimized browsing history dataset of 3 million citizen, uncovering their medical information and their sexual preferences. Also, in Australia, the Department of Health publicly released de-identified medical records for 10% of the population only for researchers and it was reidentified in 6 weeks. Again, researchers were able to uniquely identify individuals in anonymized taxi trajectories, bike sharing trips, subway data, and mobile phone and credit card datasets.

So, AI may increase the scope of application of GDPR because it can be extended to the text that are only apparently anonymous but they can be identified by ML algorithms. This problem can be addressed by:

1. Ensuring that data is deidentified in ways that make it more difficult to reidentify it;
2. Implementing security processes and measures for the release of data.

Inferred personal data

AI systems may infer new information about data subjects. The issue, in this case, is whether the inferred information should be considered as new personal data. For example, if an individual sexual orientation is inferred from their facial features or the persolnality type is inferred from online activity, is this information personal data?

If the inferred information counts as new personal data, this would trigger all the consequences that the processing of personal data entails according to the GDPR unless there is another ground that justifies their processing.

Profiling

> ‘profiling’ means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person’s performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.

Profiling is applied in social networks to make personalized advertisements or to show us the content that is more related to our interests. Profiling is also applied on predicting the likelihood of heart disease of applicants for insurance, the creditworthiness of loan applicants, the likelihood that convicted persons may reoffend and etcetera.

Profiling is also connected to the issue of automated decision making because the prediction may lead to a decision, which can be automatic or human based on the prediction.

AI and Big Data, in combination with the availability of extensive computer resources, have greately increased the opportunities for profiling. In fact, a trained ML system, given predictors-values concerning a new individual, is able to infer a corresponding target value for that individual.

Then, when AI learns a correlation concerning a person propensity to respond to a certain stimuli, this enables the transition from prediction to behavioural modification which can be a legitimate influence or even an illegal or unethical manipulation.

Legal status of inferred data

We also need to distingish the general correlations that are captured by the learned model and the results of applying that model to a particular individual. In fact, considering an ML model trained on personal data, once the model has finished the learning it doesn't contain personal data anymore, since it links any possible combination of possible input values to a corresponding likelihood and these correlations apply to all individuals sharing similar characteristics. But, then, when the model is applied to a new individual, the description of the individual and the outcome of the model represent personal data, the former being collected data and the latter being inferred data.

Being personal data, data protection rights should in principle apply and, for example, according to the Article 29 Working Party (now called Data Protection Board, a body established at European level that includes the representatives of Data Protection authorities of all European countries and they issue very importan opinions although not legally binding), in case of automated inferences, data subjects have the right to access both the personal data used as input for the inference and the one obtained as inferred output.

According to the Data Protection Regulation, there is also the right to rectification, that is when a personal piece of data is wrong the individual could ask to fix it. An interesting issue is whether a rectification only applies to the collected data or it also applies to the inferred data. For example, is it possible to ask for the rectification of the outcome of a certain prediction system? A convincing opinion is that when the inference is statistically inappropriate then you have the right to ask for a correction of the inference that has been done at least if you are able to show either that it is incorrect based on the data available either there are additional data concerning yourself showing that, by including them, a different output would have been presented.

Some lawyers have been arguing that automated inference should respect some standards and statisfy the following criteria:

Acceptability: the input data should be acceptable and not based to prohibited features (e.g sexual orientation, race);

Relevance: the inferred information should be relevant to the purpose of the decision (e.g ethnicity should not be inferred for the purpose of giving a loan);

Reliability: the input data and the methods used to process them should be reliable.

Consent

> ‘consent’ of the data subject means any freely given, specific, informed and unambiguous indication of the data subject’s wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her.

According to GDPR consent should be freely given, specific, informed, unambiguous and be expressed through a clear affermative action. In fact, there are discussions on whether or not the default acceptance setting of certain websites on the usage of cookies and personal information is really a consent in which you express the intention to allow the processing of your data through an affirmative action. Another issue may be if a consent is really free when you have to choose whether or not to consent the processing of data to be able to use a specific service.

There have been some criticisms concerning consent, even though they don't regard AI:

1. Usually users consent even if they don't have a knowledge of the processing at stake nor a real opportunity to choose;
2. At the time of consent, it is not included the future, often unknown, use of the data and so users are not aware of that.

Consent according to the GDPR must follow some rules:

- **Specificity**: consent needs to be specific, so that it cannot extends beyond what is explicitly indicated. You must know the purpose for which the data has to be processed and consent to that particular activity and the fact that the data subject has only consented to the processing for a certain purpose does not necessarly rule out that the data can be processed for a further legitimate purpose. This requirement of specificity is attenuated for scientific research which allows consent to be given not only for specific research projects but also for areas of scientific research because it is often not possible to fully identify the purpose of personal data processing for scientific research at the time of data collection;
- **Granularity**: there shoud be a consent for each separate activity;
- **Freedom**: consent should not provide a valid legal ground for the processing of personal data in specific case where there is a clear imbalance between the data subject and the controller. That's the case when a party has a market dominance

or consent is required by the provider of a service even though the processing is not necessary for performing the service.

AI and Data protection principles

These are broad principles useful for lawyers and also for computer scientists that have to keep them in mind when developing an application that concerns personal data.

Transparency

Users of a system should know how the system is going to process their data and this information should be provided in a coincise and easily accessible way.

Fairness

Users shoud not be tricked into processings which they are not aware or they are not intended to. There is a big discussion in data protection community regarding Dark Patterns, that are various ways in which websites trick the users into accepting the processing of their data by presenting the choices in an ambigous or unclear way, making them much more difficult to refuse.

Concerning fairness of content of an autometed inference or decision, there is the substantive fairness which says that in order to ensure fair and transparent processing with respect to the data subject, the controller should use appropriate mathematical or statistical procedures for profiling, implement technical measures to ensure there are no mistakes, that data is secure and decisions are not discriminatory.

Purpose limitation

Data should be collected only for a purpose that is specified, explicit and legitimate. This has some problems with AI and Big Data because their idea is that, once you have the data, you can use it for new purposes as a resource to discover hidden patterns even if you have not an idea of what these patterns could be in advance. So there is the issue of how to reconcile the purpose with the repurpose, the possibility of using already collected data for new purposes. The idea is that some reuse of data is acceptable but only when it is not incompatible with the purpose for which the data has been collected.