DOI: 10.13166/WSGE//XUXL1897

Daniela Ježová
Comenius University, Law Faculty
European Union Law Department
Jezova.daniela@flaw.uniba.sk

ARTIFICIAL INTELLIGENCE AND PRIVACY

Abstract
At the current moment due to the enormous increase of digitalization the AI´s potential is starting to show its face. Many online applications used on daily basis (Facebook, Google, Netflix) are using AI technologies to improve their services. Robots operating on AI are becoming part of our lives. The article deals with the connection between AI and privacy. AI fuel are data and we will research if the privacy is protected even in the era of AI technologies.

Keywords: *Artificial Intelligence, Privacy, European Union*

Introduction

As a starting point of the research the definition of artificial intelligence shall be introduced. There are several definitions of AI introduced by other scientists, some US states and European Union. Bases on European Union Ethical guidelines for Trustworthy AI by High – Level Expert Group on AI it is defined as software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action to take

to achieve the given goal.[1] Based on the guidelines trustworthy AI has three components, which should be met throughout the system's entire life cycle:

1. it should be lawful, complying with all applicable laws and regulations;
2. it should be ethical, ensuring adherence to ethical principles and values;
3. it should be robust, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm.

Author Turner[2] defines AI as the ability of a non-natural entity to make choices by an evaluative process.

For a law of AI a key factor is an agent, described by scientists[3] as anything that can be viewed as perceiving its environment through sensors as acting upon the environment through effectors. A robotic agent substitutes cameras and infrared range finders for the sensors and various motors for the effectors. The job of AI is to design the agent program: a function that implements the agent mapping from precepts to actions. This program will run on computing device and is called architecture. They use such a relationship description: agent = architecture + program. Another key term for AI is needed and that is autonomous. An agent is autonomous to the extent that its action choices depend on its own experience, rather than on knowledge of the environment that has been built-in by the designer. This is the ability to learn to compensate for partial or incorrect prior knowledge.

Artificial intelligence (AI) is the concept used to describe computer systems that are able to learn from their own experiences and solve complex problems in different situations – abilities we previously thought were unique to mankind. And it is data, in many cases personal data, that fuels these systems, enabling them to learn and become intelligent. Nowadays AI starts to be used by consumers on daily basis in their households and becomes a part of our lives. For example an intelligent chatbot (a computer program that people can interact with by means of ordinary speech, or through written input) analyses all the information it is fed – a combination of questions posed by customers and responses communicated by customer service.

---

[1] Ethics Guidelines For Trustworthy AI, High-Level Expert Group on Artificial Intelligence, April 2019, European Commission, Brussels available https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (06.06.2021)

[2] Turner, J.: *Robot Rules, Regulating Artificial Intelligence*, Plagrave macmillan, 2019, p. 16

[3] Russel, S., J., Norvig, P.: *Artificial Intelligence: A Modern Approach*, Prentice Hall, New Jersey, 1995, p. 31, 35

There are two main aspects of artificial intelligence that are of particular relevance for privacy. The first is that the software itself can make decisions, and the second is that the system develops by learning from experience. For a computer system to learn, it needs experience, and it obtains this experience from the information that we feed into it. This input may be in several different formats.

With the broader daily usage of AI, challenges to the current data protection law are arising. Article 8 of European Charter of Fundamental Rights provides the general basis of protection of personal data which is the basic stone of the later General Data Protection Regulation (GDPR). The privacy legislative does not address artificial intelligence in name but refer to "automated decisions or systems".

AI and GDPR

Our world is undergoing an information Big Bang, in which the universe of data doubles every two years and quintillions of bytes of data are generated every day.[4] As artificial intelligence evolves, it magnifies the ability to use personal information in ways that can intrude on privacy interests by raising analysis of personal information to new levels of power and speed. AI is not explicitly mentioned in the GPDR, but many provisions in the GDPR are relevant to AI, and some are indeed challenged by the new ways of processing personal data that are enabled by AI. There is indeed a tension between the traditional data protection principles – purpose limitation, data minimisation, the special treatment of 'sensitive data', the limitation on automated decisions– and the full deployment of the power of AI and big data. The GDPR allows for the development of AI and big data applications that successfully balance data protection and other social and economic interests, but it provides limited guidance on how to achieve this goal. It indeed abounds in vague clauses and open standards, the application of which often requires balancing competing interests.

[4] Marr, B., "*How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read*," Forbes, May 21, 2018.

Purpose limitation principle

The purpose limitation principle (art. 5 and 6 GDPR) means that the reason for processing personal data must be clearly established and indicated when the data is collected. This is essential if the data subject is to exercise control over the use of his/her information. The purpose of the processing also needs to be fully explained to the data subject if he or she is to be able to make an informed choice about whether to consent to it. This may be difficult to achieve in Big Data scenarios. At the time personal data is collected, it may still be unclear for what purpose it will later be used. However, the blunt statement that the data is collected for (any possible) Big Data analytics is not a sufficiently specified purpose. Yet the development and application of artificial intelligence often requires many different types of personal data – information that in some cases has been collected for other purposes. For example, it is possible that a person's social media activities are built into an algorithm that determines whether she will obtain a mortgage from the bank. Such recycling of information may be useful and provide more precise analyses than those which were technically feasible previously, but it does not comply with the purpose limitation principle. In cases where previously retrieved personal data is to be re-used, the controller must consider whether the new purpose is compatible with the original one. Otherwise new consent is required or the basis for processing must be changed. In the example discussed above, the data subject must consent to social media information being used by the bank in connection with mortgage applications.

The purpose limitation principle is highly important in ensuring that the data subject exercises control over his or her own personal information. Exceptions to the principle are available such as, taking place in connection with scientific or historical research, or for statistical purposes. In connection with recital 159 based on which the scientific research should be interpreted broadly and include technological development and demonstration, basic research, as well as applied and privately financed research it means that in some cases the development of may be considered to constitute scientific research. The use of personal data for scientific research is governed by specific rules in the GDPR (Article 89). Use in such instances must be subject to the appropriate safeguards to secure the data subject's rights and freedoms. The safeguards

must ensure that technical and organisational measures are in place to protect the data minimisation principle.

The purpose limitation principle is associated with number of other regulations and principles such as the duty to provide information based on art. 13 and 14 GDPR and the principle of necessity (art. 5 para 1c) GDPR). Thus, the conventional objective of the current data protection law rationalises data processing by only allowing the personal data to be processed on a statutory basis, for specific purposes and in a transparent manner. The entire approach is guided by the idea that courses of action and decision-making processes could be almost completely foreseen, planned, and steered by legal means.[5]

A tension exists between the use of AI and big data technologies and the purpose limitation requirement. These technologies enable the useful reuse of personal data for new purposes that are different from those for which the data were originally collected. To establish whether the repurposing of data is legitimate, we need to determine whether a new purpose is 'compatible' or 'not incompatible' with the purpose for which the data were originally collected. According to the Article 29WP, the relevant criteria are (a) the distance between the new purpose and the original purpose, (b) the alignment of the new purpose with the data subjects' expectations, the nature of the data and their impact on the data subjects' interests, and (c) the safeguards adopted by the controller to ensure fair processing and prevent undue impacts.[6]

The type of analytics application used can lead to results that are inaccurate, discriminatory, or otherwise illegitimate. An algorithm might spot a correlation, and then draw a statistical inference that is, when applied to inform marketing or other decisions, unfair and discriminatory. This may perpetuate existing prejudices and stereotypes and aggravate the problems of social exclusion and stratification. Further, and more broadly, the availability of large datasets and sophisticated analytics tools used to examine these datasets may also increase the economic imbalance between large corporations on one hand and consumers on the other. This economic imbalance may lead to unfair

[5] Albers, M.: *Realizing the complexity of data protection*, In.: Gutwirth, S., Leenes, R., De Hert, P., (eds) Reloading the complexity of data protection: multidisciplinary insights and contemporary challenges. Springer, 2014, p. 213

[6] Article 29 Data Protection Working Party, Opinion 03/2013 on purpose limitation, 02 April 2013

price discrimination with regard to the products and services offered, as well as highly intrusive, disruptive, and personalised targeted advertisements and offers. It could also result in other significant adverse impacts on individuals, for example, with regard to employment opportunities, bank loans, or health insurance options.

To identify what safeguards are necessary, it may be helpful to make a distinction between two different scenarios. In the first one, the organisations processing the data want to detect trends and correlations in the information. In the second one, the organisations are interested in individuals. In the first scenario, the concept of functional separation plays a key role, and the extent to which this may be achieved could be an important factor in deciding whether further use of the data for (marketing or other) research can be considered compatible. In these cases, data controllers need to guarantee the confidentiality and security of the data and take all necessary technical and organisational measures to ensure functional separation. The second potential scenario is when an organisation specifically wants to analyse or predict the personal preferences, behaviour and attitudes of individual customers, which will subsequently inform 'measures or decisions' that are taken with regard to those customers. In these cases, free, specific, informed, and unambiguous 'opt-in' consent would almost always be required, otherwise further use cannot be considered compatible. Importantly, such consent should be required, for example, for tracking and profiling for purposes of direct marketing, behavioural advertisement, data-brokering, location-based advertising, or tracking-based digital market research.

For the consent to be informed, and to ensure transparency, data subjects/consumers should be given access to their 'profiles', as well as to the logic of the algorithm that led to the development of the profile – organisations should disclose their decisional criteria. Further, in many situations, safeguards such as allowing data subjects/customers to have direct access to their data in a portable, user-friendly and machine-readable format may help empower them, and redress the economic imbalance between large corporations on one hand and data subjects/consumers on the other. It would also let individuals 'share

the wealth' created by big data and incentivise developers to offer additional features and applications to their users.[7]

Today, many data controllers use consent to legitimize data processing for purposes that would otherwise not be lawful as they exceed the initial purpose. However, using consent to legitimize otherwise illegitimate data processing has been criticised. Consumer rights organisations have pointed out that it allows data controllers to circumvent purpose limitation and makes it hard for consumers to understand contemporary data flows.[8] Indeed, many individuals seem unaware of the kinds of data processed, such as what has been termed as 'bastard data': where the merging and comparing of data results in additional personal data.[9] The GDPR defines consent as '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'. As a consequence, some forms of 'consent' such as pre-ticked boxes do not meet the GDPR threshold. In fact, research conducted in the United Kingdom in 2019 revealed that whereas 63% accept that online services are funded by advertisements, acceptance rates shift radically to only 36% once it is explained that personal data beyond browsing history is used to personalise adverts.[10] This indicates that if consent really were 'informed' most users would often not consent.[11]

---

[7] Article 29 Data Protection Working Party, Opinion 03/2013 on purpose limitation, 02 April 2013, Annex 2

[8] Biega, A., Finck, M.: Reviving Purpose Limitation and Data Minimalisation in Personalisation, Profiling and Decision-Making Systems, Max Planck Institute for Innovation and Competition Research Paper No. 21 – 04 https://arxiv.org/ftp/arxiv/papers/2101/2101.06203.pdf (10.06.2021)

[9] 'ENDitorial: Is "Privacy" Still Relevant in a World of Bastard Data?' (European Digital Rights (EDRi)) https://edri.org/our-work/enditorial-is-privacy-still-relevant-in-a-world-of-bastard-data/ (10.06.2021)

[10] Information Commissioner's Office, 'AdtechMarket Research Report' (March 2019) 5, 19 https://ico.org.uk/media/about-the-ico/documents/2614568/ico-ofcom-adtech-research-20190320.pdf (10.06.2021)

[11] Cate, FH., Cullen P., Mayer-Schonberger, V.: *Data protection principles for 21st century: revising the 1980 OECD guidelines*, 2014, available https://www.oii.ox.ac.uk/archive/downloads/publications/Data_Protection_Principles_for_the_21st_Century.pdf (06.06.2021)

Data minimalization principle

On the other hand, the principle of data minimisation requires that the data used shall be adequate, relevant, and limited to what is necessary for achieving the purpose for which the data is processed. This means that a controller cannot use more personal data than is necessary, and that the information selected must be relevant to the purpose. In some circumstances, adequacy will have a limiting effect on the quantity of data to be processed, such as where data that is inadequate in light of the purpose for which it is collected. However, in other circumstances, adequacy may require the processing of more data for data analysis to be fair and accurate.[12]

A challenge when developing AI is that it may be difficult to define the purpose of processing because it is not possible to predict what the algorithm will learn. The purpose may also be changed as the machine learns and develops. This challenges the data minimisation principle as it is difficult to define which data is necessary…However, data minimisation is more than a principle limiting the amount of detail included in training or in the use of a model. The principle also stipulates proportionality, which restricts the extent of the intervention in a data subject's privacy that the use of personal data can involve. This may be achieved by making it difficult to identify the individuals contained in the basic data. Minimisation does not exclude the inclusion of additional personal data in a processing, as long as the addition of such data provides a benefit, relatively to the purposes of the processing that outweigh the additional risks for the data subjects. Even the utility of future processing may justify retaining the data, as long as adequate security measures are in place.

The degree of identification is restricted by both the amount and the nature of the information used, as some details reveal more about a person than others. The use of pseudonymisation or encryption techniques protect the data subject's identity and help limit risks and increase therefore the compatibility of retention with minimisation.

---

[12] Biega, A., Finck, M.: Reviving Purpose Limitation and Data Minimalisation in Personalisation, Profiling and Decision-Making Systems, Max Planck Institute for Innovation and Competition Research Paper No. 21 – 04 https://arxiv.org/ftp/arxiv/papers/2101/2101.06203.pdf (10.06.2021)

The processing of personal data for merely statistical purposes may be subject to looser minimisation requirements. In such a case the data subjects' information is considered only as an input to a training set (or a statistical database) and is not used for predictions or decisions concerning individuals, which is stated in Recital 162 GDPR. Since the data subject is not individually affected by statistical processing, the proportionality assessment, as far as data protection is concerned, concerns the comparison between the (legitimate) interest in obtaining the statistical results, and the risks of the data being misused for non-statistical purposes. The information used to ascribe a person to a group and the person's ascription to that group are personal data, and so are the consequentially inferred data concerning that person. This idea is expressed in at footnote 5 in the 2017 Council of Europe Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data.

This principle also forces developers to thoroughly examine the intended area of application of the model to facilitate selection of relevant data necessary for the purpose. Furthermore, the developer must consider how to achieve the objective in a way that is least invasive for the data subjects. The assessments performed need to be documented, so that they can be presented to the Data Protection Authority in the event of an inspection, or in connection with a preliminary discussion…The principle of data minimisation should play a major role in the development of artificial intelligence so that the rights of data subjects are protected and general confidence in the models retained.

Thanks to AI and big data, it may be possible to link observable behaviour and known features of individuals – online activity, purchases, likes, movements – to non-observable sensitive data on them such as their psychological attitudes, their health condition their sexual orientation, or their political preferences. Such inferences may expose the concerned individuals to discrimination or manipulation.

Many agree that data minimisation not only entails an obligation to restrict the amount of data but also to keep sensitive data to a minimum.

Transparency

The self-adaptive "behaviour" of at least certain types of AI technologies leads to a lack of transparency. This phenomenon is often referred to as the black box issue of AI technologies. Why is this a problem for the traditional approach to evaluating data protection?

Transparency is achieved by providing data subjects with process details. Data subjects must be informed about how the information will be used, whether this information is collected by the data subjects themselves or by others (GDPR Articles 13 and 14). Besides, the information must be easily available, on a home page for example, and be written in a clear and comprehensible language (GDPR Articles 12). This information shall enable the data subjects to exercise their rights pursuant to the GDPR. The idea of transparency is specified in Recital 58, which focuses on conciseness, accessibility and understandability. Information must also be provided about 'the legitimate interests pursued by the controller or by a third party' where the processing is based on legitimate interest (Article 6(1)(f)). When the data are processed for purposes that could not be foreseen at the time the data were collected – as it is often the case with machine learning applications– the information has to be provided before the new processing. The obligation to inform the data subject is waved when compliance is impossible, requires a disproportionate effort or impairs the achievement of the objective of the processing.

It can be challenging to satisfy the transparency principle in the development and use of artificial intelligence. Firstly, this is because the advanced technology employed is difficult to understand and explain, and secondly because the black box makes it practically impossible to explain how information is correlated and weighted in a specific process. It is also challenging that information about the model may reveal commercial secrets and intellectual property rights, which according to the GDPR's preface (Recital 63) the right of access must avoid. Furnishing the data subject with the information she needs to protect her interests, without at the same time disclosing trade secrets, will not be problematical.

Automated decision-making

Article 22, which deals with automated decision-making, is most relevant to AI. As we shall see in what follows, this provision combines a general prohibition on automated decision-making, with broad exceptions.

Individual automated decisions are decisions relating to individuals that are based on machine processing. An example of this is the imposition of a fine on the basis of an image recorded by an automatic speed camera. Automated decisions are defined and regulated in Article 22 of the GDPR. Essentially, automated individual decisions are not permitted. Exceptions apply, however, if the automated decision is a necessary condition for entering into a contract, is permitted by law, or is based on the explicit consent of the data subject. In order to meet the requirements of the Regulation, the decision must be based solely on automated processing, includes profiling and it must produce legal effect, or similarly significantly affect a person. There cannot be any form of human intervention in the decision-making process.

Article 22 provision refers to a right, it does not provide for a right to object to automated decision-making, namely, it does not assume that automated decision-making is in general permissible as long as the data subject does not object to it.

Many decisions made today by AI systems fall under the scope of Article 21(1), as AI algorithms are increasingly deployed in recruitment, lending, access to insurance, health services, social security, education, etc. The use of AI makes it more likely that a decision will be based 'solely' on automated processing.

Conclusion

In conclusion, while AI is a pioneering and rapidly developing technology, it does fit in with the existing legal framework for data protection. The GDPR was introduced with an intention to guarantee lawful and fair processing of personal information in an age where data is used as a commodity and resource. Through the introduction of several core principles, it manages to create a uniform system of safeguards for data subjects, which can be applied regardless of the used technology or mechanisms. This conclusion extends to artificial intelligence as well, the use of which can be easily exploited once the provisions and principles of the GDPR are sufficiently understood and adhered to. As has been presented above AI comes in conflict with principles of data protection or at least a certain tension is between them. The tension with the principle of transparency, the purpose limitation, data minimalization which oppose the idea of big data to collect as many data as possible for the purpose of pattern recognition. Purpose limitation does not have the capacity to restrict data collection and analysis. It is explicit, legitimate and formulated with enough specificity, any purpose is a valid purpose under the GDPR. Service providers can always choose to go beyond the predefined purpose and process more data as soon as they get data subject consent. Data subjects are likely to provide such consent as they fail to understand implications thereof and want or need to continue using a given service. The above analysis has confirmed that data minimisation continues to play an important role in data processing systems as it compels service providers to think systematically about what data they need to achieve a given purpose. However, from a computational perspective, data minimisation requires the development of specific decision rules and mathematical measures defining when individual pieces or sets of data are 'adequate', 'necessary', and 'relevant'.