Polish Political Science Yearbook, **vol. 51(3) (2022), pp. 183–191**
DOI: https://doi.org/10.15804/ppsy202240 PL ISSN 0208-7375
www.czasopisma.marszalek.com.pl/10-15804/ppsy

Katarzyna Chałubińska-Jentkiewicz
War Studies Academy (Poland)
ORCID: 0000-0003-0188-5704
e-mail: k.jentkiewicz@akademia.mil.pl

Monika Nowikowska
War Studies Academy (Poland)
ORCID: 0000-0001-5166-8375
e-mail: m.nowikowska@akademia.mil.pl

Artificial Intelligence v. Personal Data

Abstract: The world is constantly changing under the influence of new technologies. Artificial intelligence systems are currently used in many areas of human activity. Such systems are increasingly assigned the tasks of collecting and analysing personal data. The areas successfully using AI include transport, medicine, trade, marketing, and others. The number of these areas increases proportionally with the advancement of technology. We can process vast amounts of data and analyse it using IA. It is, of course, big data that sits at the heart of AI. As computing systems generally have grown in power and capacity, data consumption has grown exponentially.

Keywords: *Artificial Intelligence (AI), new technologies, personal data, regulation*

Introduction

Personal data protection is one of the most sensitive challenges faced by contemporary legal science, arising from the extraordinary technological advancement that occurred in the last decades. The article is devoted to the issue of AI v. Personal Data. The development of new technologies creates new challenges and threats. One of the most interesting examples is the creation of AI systems focused on permanent self-development. For this purpose, they use various data types, including personal data. The regulation of AI is one of the significant challenges faced by the EU (Schreiber, 2020). Most researchers focus on the substantive scope of AI regulation, including state law, soft law and ethical norms. When we add AI, big data and machine learning into that equation, it is clear that we must be ready for change. Companies that work with these technologies will get a competitive edge, and those that ignore them, no matter what industry they operate in, will face the risk of extinction (Buyers, 2018).

However, the issue of new technologies also means threats. In an increasing number of portals, we are greeted by virtual assistants who obtain information about us. For private individuals, the greatest threat may be the automation of people's work by AI mechanisms. It is about so-called phishing, i.e., extracting confidential data from users on the web. So far, it has been done by another human being. Today, AI is so reliable that it does not need people. The advantage of artificial intelligence is that it can conduct millions of such conversations simultaneously using chats or e-mails. AI successfully steals data while it seems we are talking to a human.

There are many concerns and questions around AI, but a more pressing concern relevant to today is the protection of personal data (Marszałek-Kawa & Plecka, 2019). These technologies are becoming widespread, and the amount of data collected is increasing in variety and volume. Data collected from various sources is being used to develop AI using machine learning and deep learning algorithms, and as a result, data has become the food for AI.

The issues raised required an analysis of the content and evaluation of the literature on the subject (the application of the desk research technique) and selected acts of EU and Polish law, covering three basic issues: the concept of personal data, artificial intelligence – application and perspectives, AI subject to the General Data Protection Regulation (GDPR). It is important to develop technologies under the law, GDPR, and democracy (Marszałek-Kawa, 2019).

Legal Basis

On February 19, 2020, the European Commission published a „White Paper on Artificial Intelligence – A European approach to excellence and trust (the White Paper) (Brussels, 19.2.2020 COM(2020) 65 final). The document underlines that AI is developing fast. It will change our lives by improving healthcare, increasing farming efficiency, contributing to climate change mitigation and adaptation, increasing the security of Europeans, and in many other ways that we can only begin to imagine. At the same time, AI entails many potential risks, such as opaque decision-making, gender-based or other kinds of discrimination, intrusion into our private lives or being used for criminal purposes. Against a background of fierce global competition, a solid European approach is needed, building on the European strategy for AI. To address the opportunities and challenges of AI, the EU must act as one and define its own way, based on European values, to promote the development and deployment of AI.

As digital technology becomes an ever more central part of every aspect of people's lives, people should be able to trust it. Trustworthiness is also a prerequisite for its uptake. It is a chance for Europe, given its strong attachment to values and the rule of law and its proven capacity to build safe, reliable, and sophisticated products and services. However, do we trust this computer? „Trust this computer?" is the question that iTunes asks when a user connects an iPhone to a PC before initiating data transfer. It is a very good question because

„trust" is the fundamental feeling behind entrusting someone with something that belongs to you. But how can an individual trust a system or AI? We believe the best way to create this trust is to establish clear regulations protecting personal data (Sobczak et al., 2022).

Data protection principles and rules are not barriers against AI and big data and can be used by people working in these fields as guidance to increase data security, data quality and data volume/variety by gaining the trust of individuals. Therefore, a legislative framework on personal data is a must for the controlled development of these technologies.

Poland, with a population of almost 40,000,000 and over 28,000,000 internet users, is a big personal data market. Many people work on these technologies, which require a legal framework for processing personal data. There is no specific legislation on big data or AI in Poland. Most informed citizens probably know by now that corporations, using AI, collect information about them, but they may well be unaware of the extent and scope of widespread invasions of privacy. Many may be away from tracking tools called "cookies", which are installed on one's computer by websites. They are used to identify the person and to remember his or her preferences (Etzioni, 2015).

The principles of personal data protection have been regulated under Polish law in several legal acts. The fundamental act which stipulates the protection of personal data is the Constitution of the Republic of Poland of 2 April 1997. The right to personal data protection is a unique legal construct intended to protect the values referred to in Article 47 of the Constitution of the Republic of Poland. The Constitution provides that everyone is entitled to the legal protection of their private life, family life, honour, and reputation, as well as the right to decide on their personal life. In the relevant literature, the individual's right to protect their personal data is called "information autonomy". The right to the protection of personal data is categorically associated with the right to privacy, recognising it as its unique form (Nowikowska, 2018).

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 the free movement of such data, and repealing Directive 95/46/EC (The General Data Protection Regulation) (GDPR), is also of fundamental importance in this regard. The issue of personal data is also governed by the Act of 10 May 2018 on personal data protection, which repealed several provisions of the former Act, and introduced new ones, which regulate, inter alia, the status of the President of the Personal Data Protection Office, as well as the procedure for initiating and conducting proceedings in connection with the infringement of personal data in the common courts, and the Act. The group of legislative acts regulating the principles of personal data processing in cyberspace also includes the Act on the National Cybersecurity System.

According to Article 1 of the GDPR, the EU legislators, when determining the adoption and application of uniform solutions for the processing of personal data in all EU Member States, pursue two equally important objectives: first, they protect the fundamental rights and freedoms of natural persons, and in particular, the right to the protection of their

personal data; and second, they ensure the free transfer of personal data between Member States (Nowikowska, 2021).

Personal data under GDPR is regarded as any personal information relating to a natural person. Whether private, professional, or part of public life, it is considered personal data. This data includes anything that could be used to identify an individual, either directly or indirectly, even where that data is considered generalised (Gobeo et al., 2018). Some categories are a person's name, identification numbers such as a social insurance number, ID number, age, location, physical, mental, genetic, sexual orientation, medical records, email, social, and more (Walters et al., 2019). It should be emphasised that no comprehensive list of categories is given within the GDPR. The broad range of data sources becomes especially relevant where profiling of data is used. Personal data covering all these categories can provide comprehensive insight into a given individual (Voigt & Bussche, 2017).

In European Union, the GDPR protects personal data regardless of the technology used for processing that data – it is technology neutral and applies to both automated and manual processing, provided the data is organised under pre-defined criteria. It also does not matter how the data is stored – in an IT system, through video surveillance, or on paper. In all cases, personal data is subject to the protection requirements set out in the GDPR (Chałubińska-Jentkiewicz & Nowikowska, 2021).

There is an overused refrain, which is often heard when data protection is broached: "I have nothing to hide, there is nothing interesting about me". Alternatively: "If you have nothing to hide, then you have nothing to fear". That implies that only "bad people", criminals or terrorists have something to fear from the analysis and the exposure of their private information. It also implies that the only people who can justify the desire for privacy are those same "bad people" (Gobeo et al., 2018).

Ultimately, privacy is sacrificed by the individual in exchange for not being labelled as "bad". The dangers inherent in surveillance, through collecting and using personal data, are mitigated only when the individual behaves unthreateningly. It is constructed as complying with the status quo, or at the very least, to keep dissent hidden and unspoken. It applies equally in politics, commerce, and the social sphere.

Artificial Intelligence and Privacy – Issues and Challenges

This article serves as an introduction to a broader conversation regarding information privacy and AI. This resource aims to: provide a high-level understanding of AI and its uses in the public sector and highlight some of the challenges and opportunities that AI presents in relation to information privacy.

In October 2018, a Personal Data Protection Commissioners conference was held in Brussels. The assumption was to establish rules regarding GDPR and artificial intelligence. They include (1) transparency, (2) purpose limitation, (3) data minimisation, (4) accountability, (5) Privacy by Design. It should be remembered that artificial intelligence uses databases

and learns from their analysis. Hence, there is a need to comply with the provisions of the GDPR, including storage limitation and minimisation. In addition to basic data, AI can also process biometric data. According to the GDPR, such action is prohibited. The exception is the user's explicit consent or the legal provision that authorises the administrator. Regardless of the type of data stored, we must keep it secure and confidential. The topic of artificial intelligence is current and extremely important from the point of view of personal data protection in the context of rapid technological development, which aims to revolutionise the reality that surrounds us.

This issue has recently been brought to attention by the British data protection authority – Information Commissioner's Office (ICO). Its guide indicates the most important issues linking artificial intelligence systems with the law on the protection of personal data and is also an attempt to find an answer to the question of how to safely and legally combine these two sometimes contradictory issues. The ICO guide indicates the need for data minimisation and privacy protection techniques when creating or introducing AI systems for their activities. The British authority emphasises that the first contradiction between the creation of artificial intelligence systems and the law on the protection of personal data is the scope of the collected data. As indicated by the authority, AI systems need huge amounts of data for proper operation and development, while the GDPR rules on the adequacy and minimising the scope of data processed. Therefore, at the very beginning, data controllers are faced with difficult requirements because it is their responsibility to demonstrate the necessity of a specific scope of data. Compliance with these principles must be case-specific. However, AI developers may use certain known techniques when developing their systems (Jakubik & Świętnicki, 2020).

The five principles of the GDPR are the guiding ethical intentions underpinning the legislation. The principles set forth the vision for data protection going forward.

(1) Transparency

The GDPR requires that all personal data processing be done in a lawful, fair, and transparent matter. It means that the legality of the basis, that is the reason for the processing, must be stated and documented clearly and transparently, and to which the data controller will be held accountable. Our current understanding of information privacy rests on the ability of individuals to exercise choices regarding the information others have about them and what is done with it. However, AI's complexity can mean that processes are unclear to individuals whose information is being used, making truly informed and meaningful consent unattainable. For instance, deep learning techniques can pose challenges to transparency, as providing an explanation about how conclusions are drawn can sometimes be difficult, even for those initially developing the algorithms, let alone the average individual. Organisations will struggle to be transparent in their AI practices or to obtain consent if they cannot communicate the processes to citizens. This tripartite principle is extremely important as

it underpins all the other principles and rights within the GDPR: lawfulness, transparency and fairness (Gobeo et al., 2018).

(2) Purpose Limitation

The GDPR clarifies the circumstances under which personal data may be processed by highlighting areas that have often been ignored. The collection of personal information should be limited to only what is necessary. Personal information should only be collected by lawful and fair means and, where appropriate, should be collected with the knowledge or consent of the individual. Personal data must only be processed for the original purposes for which it was collected. It must be collected for a specific and explicit purpose, and that purpose must have a lawful basis which forms the only reason upon which that data may be processed (Gobeo et al., 2018). The purpose of collecting personal information should be specified to the individual at the time of collection. It means that personal information should only be used or disclosed for the purpose it was collected unless there is consent or legal authority to do otherwise. The underlying goal of these intertwined principles is to minimise the amount of information an organisation holds about an individual and to ensure that the way the information is handled is consistent with the expectations of that individual.

The ability of AI to extract meaning from data beyond what it was initially collected for presents a significant challenge to this principle. In some cases, organisations may not know how AI will use the information in the future. There is a risk of excessive data collection beyond what is necessary 'just in case', using overly broad collection notices. This kind of practice allows organisations to claim technical compliance with their privacy obligations, but it is disingenuous and inconsistent with the underlying goal of the collection limitation principle. Further, it undermines the ability of individuals to exercise meaningful control over their personal information.

AI's basic task is to develop and improve adequate action planned by the creator. The functions of artificial intelligence and machine learning may cause the processing of personal data to take place in various ways and are also used for purposes other than those for which the solutions were originally programmed. It may result in a complete loss of control of personal information. The issue of control and audit is also important. It may prove difficult to carry out controls and audits. The complexity of some algorithms that are the building blocks of AI solutions will require research with the participation of competent people. Artificial intelligence creators will also need to maintain their know-how and trade secrets, which may hinder the commonly accepted practice of periodic system operation checks.

(3) Data Minimisation

The personal data collected should be only that data which is adequate, relevant and limited to what is necessary for the specific purpose stated. Governance and oversight are championed in information privacy law to ensure appropriate structures are in place that prevent a power imbalance between citizens and government. Minimisation means collecting only the data that can be used for specified purposes. Many businesses collect far more than is required, believing that it may become useful one day. The GDPR requires controllers to demonstrate the need for collecting and processing personal data and where that need cannot be demonstrated as non-compliant (Gobeo et al., 2018).

Machine learning is the most commonly used artificial intelligence operation. People responsible in organisations for risk management and compliance of AI systems should be aware of such techniques and be able to implement appropriate solutions with IT departments. The default approach of data scientists in designing and building AI systems does not necessarily consider the constraints of data minimisation. Organisations, therefore, need to implement appropriate risk management practices to ensure that, by design, data minimisation requirements and all relevant minimisation techniques are fully taken into account (Jakubik & Świętnicki, 2020).

(4) Accountability

Whereas the former Data Protection Directive did not explicitly emphasise accountability, the GDPR introduces the general principle of accountability in art. 5 sec. 2. This article imposes the responsibility for the compliance of processing with the GDPR and the burden of proof for said compliance onto the controller. Thus, the principle of accountability consists of two elements: 1) the responsibility of the controller to ensure compliance with the GDPR and 2) the controller's ability to prove compliance to Supervisory Authorities (Voigt & Bussche, 2017).

The creation, use and maintenance of internal policies and procedures designed to assist in organisational compliance is a key tool in achieving and demonstrating the accountability principle. These policies and procedures aim to protect personal data and minimise the risk of data breaches. The GDPR describes the rights afforded to the individual in the new era of data that can also be used in the AI environment.

(5) Privacy by Design

"Privacy by Design" means "data protection through technology design". It means that data protection in data processing procedures is best respected when it is already integrated with the technology at the time of creation. The legislation leaves it open to what exact protective measures should be taken. For example, we can indicate encryption and anonymisation of

data as possible protective measures, and user authentication. When selecting individual cases, one must ensure that state-of-the-art and reasonable implementation costs are included. The concept of Privacy by Design (art. 25 sec. 1 GDPR) is based on the realisation that the conditions for data processing are fundamentally being set by the soft and hardware used for the task. When creating new technology, developers and producers shall be obliged to keep data minimisation in mind. Examples include IT systems directed towards data minimisation and comprehensive and timely pseudonymisation of personal data (Voigt & Bussche, 2017).

In addition to the named criteria, the type, scope, circumstances and purpose of the processing must be considered. It must be contrasted with the various probability of occurrence and the severity of the risks connected to the processing. The text of the law leads one to conclude that several protective measures must often be used with one another to satisfy statutory requirements. This consideration is already performed in an early development phase when setting technology decisions. Recognised certification can serve as an indicator to authorities that the persons responsible have complied with the statutory requirements of "Privacy by Design".

Artificial Intelligence in the Public Sector

While AI technology development is driven mainly by industry and academic research, AI applications and development are also relevant to the public sector. The government already uses AI in many areas, but it stands to benefit from the further adoption of these technologies. Further, the government has a significant role in shaping how AI technologies impact citizens' lives through regulation, policy, and best practices. It is important that the government is not left behind as the private sector steams ahead – this means taking a proactive, dynamic and informed approach to technology and its interaction with law and society. In the short term, AI applications have the potential to be immensely useful in increasing the efficiency of established government processes such as answering questions, filling out and searching documents, routing requests, translating, and drafting documents. As an example, the use of chatbots to provide customer service and advice to individuals already occurs in some of the larger European government organisations[1].

Concluding Remarks

Artificial intelligence methods are not perfect, but they are promising. In many areas of our lives, AI has started to appear – but is it wrong? Thanks to the existence of, among others, GDPR, we have certain rules, obligations and rights regarding personal data. Artificial

[1] https://privacyinternational.org/sites/default/files/styles/middle_column_cropped_large_1x/public/2020–06/john-noonan-QM_LE41VJJ4-unsplash.jpg?itok=9-IYrnW_

intelligence, used under the law, can be very useful in our lives – we can already see its amazing application, for example, in the medical industry. To sum up – personal data is slowly becoming a currency – this is a fact. Therefore, we should be careful about entering personal information and what we consent to on the Internet. With the careful use of artificial intelligence methods, we can calmly take advantage of its advantages. There is much research on the emergence of a „privacy paradox", in which people express concern for their privacy but, in practice, continue to willingly contribute their information via the systems and technologies they use.

References:

Buyers, J. (2018). *Artificial Intelligence. The Practical Legal Issues*. Law Brief Publishing.

Chałubińska-Jentkiewicz, K., & Nowikowska, M. (2021). *Ochrona danych osobowych w cyberprzestrzeni*. Wydawnictwo Akademii Sztuki Wojennej.

Etzoni, A. (2015). *Privacy in a Cyber Age. Policy and Practice*. Palgrave Macmillan.

Gobeo, A., Flower, C., & Buchanan, W. J. (2018). *GDPR and Cyber Security for Business Information System*. River Publishers.

Jakubik, M., & Świętnicki, T. (2020). *RODO w IT: sztuczna inteligencja a dane osobowe - czy RODO definiuje AI oraz ML? Komentarze praktyczne*. LEX/el.

Marszałek-Kawa, J. (Ed.) (2019). *Państwo w obliczu współczesnych wyzwań. O cyberbezpieczeństwie i innych zagrożeniach na przykładzie wybranych państw azjatyckich*. Wydawnictwo Adam Marszałek.

Marszałek-Kawa, J., & Plecka, D. (Eds.) (2019). *The Dictionary of Political Knowledge*. Wydawnictwo Adam Marszałek.

Nowikowska, M. (2018). Ochrona danych osobowych w dokumentach kontrolnych. In J. Taczkowska-Olszewska, M. Nowikowska, & A. Brzostek (Eds.), *Reforma ochrony danych osobowych. Cel, narzędzia, skutki*. Silva Rerum.

Nowikowska, M. (2022). Personal Data Protection in the Context of the Act on the National Cybersecurity System. In K. Chałubińska-Jentkiewicz, F. Radoniewicz, & T. Zieliński (Eds.), *Cybersecurity in Poland. Legal Aspects*. Springer.

Schreiber, A. (2020). Right to Privacy and Personal Data Protection in Brazilian Law. In D. M. Vicente, & S. Casimiro (Eds.), *Data Protection in the Internet*. Springer.

Sobczak, J., Chałubińska-Jentkiewicz, K., & Nowikowska, M. (2022). *Piractwo w sieci*. Silva Rerum.

Voigt, P., & Bussche, A. (2017). *The EU General Data Protection Regulation (GDPR). A Practical Guide*. Springer.

Walters, R., Trakman, L., & Zeller, B. (2019). *Data Protection Law. A comparative Analysis of Asia-Pacific and European Approaches*. Springer.