Medij. istraž. (god. 26, br. 2) 2020. (9-28)
IZVORNI ZNANSTVENI RAD
DOI: 10.22572/mi.26.2.1
Zaprimljeno: 2. lipnja 2020.

Towards the Trustworthy AI: Insights from the Regulations on Data Protection and Information Security

Tihomir Katulić*

SUMMARY

After decades of theoretical deliberations, the rapid development of advanced information technology has allowed machine learning as a first practical step towards artificial intelligence to enter widespread commercial and government use. The transition into a post-industrial, information society has revealed the value of data as an important resource whose processing is the basis of the new innovative information society services. The European Union has enacted several important regulations and directives in the recent past to protect the recognized fundamental rights of individuals and to regulate the obligations of service providers to ensure safe and secure processing. The Charter of Fundamental Rights as the legal basis of the European system of human rights contains significant checks and limitations to the effect and purpose of future EU AI regulation. Whenever and however this regulation is adopted, it will need to comply with and contain existing European legal standards regarding the fundamental rights of individuals in the EU. The European Commission's ethical guidelines establish ethical principles based on the recognized fundamental rights that future AI systems need to adhere to in order to be recognized as trustworthy. The purpose of this paper is to present and analyse the mechanisms present in existing European regulations in the fields of data protection and information security and in the European Union documents regarding the future artificial intelligence regulation and to offer suggestions for future regulations. The research methodology includes a comparative analysis of available regula-

* Assistant Professor Tihomir Katulić, Ph.D., Faculty of Law of the University of Zagreb, Trg Republike Hrvatske 14, Zagreb, Croatia
E-mail: tihomir.katulic@pravo.hr

tions and policy documents of the European Union, national laws, legal literature, and other sources.

Keywords: artificial intelligence, data protection, information security, information society

Introduction

The incessant march of the Moore's Law (Moore, 1965) has in recent years enabled yet another manifestation of the information revolution, fostering significant developments in the field of artificial intelligence (AI). Once a privileged field of research available to a handful of computer scientists with access to massive supercomputers with hundreds of thousands of processor cores and petabytes of working memory, the advances in parallel computing and hardware technology have democratized attempts to develop better, more efficient machine learning algorithms.

With the advent of affordable storage, processing power and broadband internet communications, a seventy-year-old promise of creating thinking machines (Solomonoff, 1985:149) has seen giant strides of technological advances in a period shorter than a decade. There is now a rapidly expanding market demand for AI services and AI technology start-ups are among the best funded and sought after by major investment industry players (Forbes, 2019). From chatbots that answer customer emails and instant messages and electronic agents that choose best airplane routs and taxi fares to the use of AI in legal and medical research artificial intelligence is starting to cause a major impact from the jobs market to the quality and availability of government services (Vozochka et al. 2018:57).

The rise of AI is having a profound effect on all aspects of science and research as well. Seemingly there are no problems or applications that cannot benefit from machine learning, intelligent agents and other manifestations of the current state of AI technology. Even at this early stage, machine learning is accelerating advances in various fields from machine translation, electronic agent intelligence, game theory and decision making to stock market trading, signal processing and medical research. (Ontañón et al., 2013:293, Dilsizian, Siegel, 2014:1, L. Chen et al. 2018:46625, etc.).

Essentially, artificial intelligence technologies are information technologies directly benefiting from vast repositories and collections of data our civilization creates as a result or a by-product of everyday activities (Sebag, 2014:11). Data in machine readable form is created, transmitted and shared globally and represents a valuable digital resource that information society exploits and monetizes in various ways, from serving personalized marketing to business intelligence and digital content

management (Malgieri, Custers, 2018:289). This resource has long since become the most valuable commodity in this post-industrial, information age, fuelling online marketing and participating in monetization of copyrighted digital content (Prins, 2016:270). Even established, successful business models in content monetization such as software licencing and software as a service are increasingly being replaced with a business model where users pay for access to content with their personal data. It has become quite obvious that leading economic blocks are developing their legal framework from a perspective influenced by political and economic issues, recently highlighted by the EU efforts to establish an effective framework for personal data protection (European Commission, 2012). This is by no means a new and revolutionary development as proven by numerous controversies around issues such as computer software copyright protection, database protection and even earlier copyright disputes going back to the end of the 19th century as the value of immaterial property became obvious for economic and social development. (Katulić, 2015:237)

Ownership of data, more to the point, ownership and rights to economic exploitation of data is becoming a dominant motive behind the recent information technology law developments.

This new legal discipline, now firmly established in comparative legal literature and legal science, is increasingly deconstructing old positivistic views on the application of traditional legal institutes (Lessig, 1999) and solutions to information society problems instead offering novel legal institutes and solutions adapted to current stage of societal development.

From Data Protection and Information Security to Digital Single Market

The European legislators have been at the forefront of this activity for well over quarter of a century with laws ranging from the Data Protection Directive in 1995 to a number of *electronic* directives in late 1990s regulating liability of information services providers, establishing rules for electronic identification services such as electronic signatures and updating the European *acquis* in the field of intellectual property. These efforts continued to include those directly applicable to the new personal data economy such as the General Data Protection Regulation, the Network and Information Security Directive as the first European law in the field of information security, the new Regulation 910/2014 on electronic identification and trust services and is an ongoing process which will soon include EU-wide regulation on class action consumer protection (European Parliament 2018), specific regu-

lation concerning digital content and service platforms as gatekeepers (Digital Services Act Package) etc.

The support for this development is both political and economic. The Barroso Commission (2005-2015) and especially Juncker Commission (2015-2020) have repeatedly underlined the necessity of enabling the European Digital Single Market to help bridge an increasing divide between EU, and the information technology leaders such as US, China, Korea and Japan, with the latter proclaiming the goal to be one of the three priorities of their tenure (EPRS 2019). Indeed, the last five years have seen an increasing amount of information society regulation coming out of Brussels, culminating with the GDPR and the NIS Directive.

The Digital Single Market was a powerful incentive. Another motive for this development is the general political tendencies of European Union and the assertion of the human rights standards contained in the Charter of the Fundamental Rights of the European Union (Butarelli, 2016:77). In an increasingly globalized world where all economy, especially data driven one, knows no territorial bounds there is going to be strife and conflict where economic and political positions are challenged (Safari, 2016:809). Protection of fundamental human rights in the digital domain has struggled with the new and innovative services and products developed on the back of the fourth information revolution.

In a recent whitepaper published by the European Commission in February of 2020, the new European Commission recognizes the numerous risks AI represents for fundamental rights recognized by the EU Charter of Fundamental Rights, such as rights to privacy and protection of personal data (Art. 7 and 8), right to freedom of expression and the right to freedom of information (Article 11), right to choose an occupation and right to work (Article 15) and so on (EU Commission Whitepaper 2020).In its Communications, while recognizing the value data driven post-industrial information society services bring to the economy of Europe, the Commission underlines the benefits AI ecosystem brings to the European society for citizens, businesses and for the public services.

Additionally, considering the state of AI in economic blocks rivalling Europe, the Commission understands that in order to stay economically and politically relevant in this century the EU needs to promote AI development and quickly close the technological gap that may turn into an economic and political one as well. An advent of AI and its mass application for citizens may result in improved health care and increased quality of life, better, more affordable, functional and higher quality products, safer and cleaner transport systems. For businesses, AI may foster development of better products and services especially in the areas where Europe already enjoys a leading position. AI will also benefit development of more efficient public services by reducing the cost in areas of transport, education, energy, waste manage-

ment and by providing tools and services to ensure a higher level of security (European Commission, 2020).

These efforts continue to have the highest level of support and attention in the EU. Von der Leyen in Political Guidelines for the Next European Commission 2019-2024 noted the importance to hold on to the European way of balancing development of new technologies while preserving highest standards of privacy, security, safety and ethical standards (von der Leyen, 2019). Game-changing technology such as AI with its many opportunities and many risks needs an effective and stable legal framework that defines the material scope of liability for all those involved with developing and implementing this technology into new innovative services and products or replacing the existing information infrastructure in services we use today. There is however a strong public opinion against such regulation, especially coming from the ranks of entrepreneurs and developers comparing the experiences of founding and financing start-up companies in Europe to the conditions in Silicon Valley or the Far East. The main argument there is that the EU legal framework is too complicated and constraining to foster efficient development of new technology and that EU administrative burden, alongside other factors, is the reason why the EU is lagging behind in innovation and, more importantly, market capitalisation of technological research. A recent statistics published by Statista in January 2020 claims that United States and China dwarf EU in the number of successful start-up companies (*unicorns,* privately owned companies with current valuation over 1 billion US dollars), with US being a home to 265 and China to 204 such companies compared to European Union's 30.

How to Regulate AI - Towards the Trustworthy AI

The problem how to regulate AI can be approached from different directions, economic, ethical and legal. The purpose of this paper is, with understanding to other approaches, to primarily consider the legal perspective with a necessary understanding of all approaches. Should EU regulate AI or not, and if yes, what should AI Regulation draw from the current state of European information technology law?

The Charter of Fundamental Rights as the modern legal basis of the European system of human rights gives very little leeway to the effect and purpose of future EU AI regulation. Whenever and however new regulation is adopted, it will need to comply and contain existing European legal standards regarding the fundamental rights of individuals in the EU. The fundamental rights are one of the crowning achievements of the European project and regardless of the current economic or political situation, these rights will have to be respected and protected. The question is then how should the European legislator go about protecting fundamental rights

while creating a legal framework for developing and adopting AI in the common market, ensuring as much as possible a level playfield with the leading blocks while not jeopardizing the rights and freedoms of individuals in the EU? Can recent experiences with regulating the use and flow of personal data and the obligations of essential and digital service providers teach a meaningful lesson applicable to a technology such as AI?

The legal literature available is just discovering the challenges of this field. While there is already a substantial amount of research conducted into various sector-specific areas of AI regulation, such as use of AI in education (Berendt, Littlejohn & Blakemore, 2019:312), AI and cybersecurity/cyberwarfare (Taddeo & Floridi, 2018), healthcare (Terry, 2019:1) etc.

Officially available documents and communications of the European Union that refer to the developments in the field of Artificial Intelligence, such as the European Commission Communications COM(2018) 237, COM (2018) 795 and COM(2019) 168) underline three basic requirements and activities – the promotion of public and private investments into AI to promote development of AI technology, the research and preparation into socio-economic changes that will be brought forth by the AI and, most importantly, ensuring that an adequate ethical and legal framework will protect and safeguard the current level of fundamental rights that individuals enjoy in the European Union (European Union Communication COM 168/2019).

The Commission calls this approach *Trustworthy AI*. In 2018 the Commission established an Independent High-level Expert Group on AI that soon delivered a document, Ethics Guidelines for Trustworthy AI*,* explaining the ethical and legal issues required to create a Trustworthy AI technology (EU High-Level Expert Group on AI, 2018). The Guidelines stipulate three principles for the artificial intelligence system which need to be implemented if the AI system is to be considered trustworthy. These principles are lawful processing, ethical operation and robust performance and all of them need to be applied together to ensure the preservation of rights and freedoms of individuals as established by the applicable laws in the EU.

The first component, lawful processing, understands that the AI system needs to perform all its data processing in a lawful manner, complying with all applicable laws and regulations. The applicable legal framework the EU has developed over decades is impressive. It regulates not only current topics and recent developments such as the new framework of personal data protection with the adoption of the General Data Protection Regulation, accompanying Directives and national implementation measures and the Network and Information Security Directive as the first European information security law, but also previous and established regulations in the fields of the freedom of information right, such as the right to access documents of the Union and reuse of public sector information, free speech, copyright etc. Ju-

dicature of European courts such as the ECJ or the ECHR that develop and protect fundamental rights on this matter is also extensive (Polcak et al, 2017).

In fact, in order to satisfy the requirements of the first component, the AI system should operate in accordance to the body of applicable European law. This includes EU primary law such as the Treaties of the European Union, the Charter of the Fundamental Rights of the European Union and the numerous secondary law sources such as Directives and Regulations in areas as diverse as data protection, information security, copyright, patents and other fields of intellectual property, consumer protection, anti-discrimination, competition, food and medicine safety, electronic commerce and electronic communications etc.

It should also account for legal obligations coming from UN Human Rights treaties, Council of Europe Conventions and applicable Member State laws. The currently applicable legal framework regulating information security in the Member States mandates application of best information security practices (NIS Directive, 2016) and has moved to the concept of assessing, preventing and mitigating information security risks that threaten information systems from a remarkably varied array of sources and actors characterized by distinct and contrasting methodology, motives and the level of knowledge and technical sophistication (Grisham et al., 2017).

In contrast to the General Data Protection Regulation, the NIS Directive was required to be transposed into national legal systems of Member States. This process took significantly longer than envisaged by the EU legislators and was finally completed in early 2020. In many respects, the information security obligations for essential service operators and digital service providers resemble the state of data protection in the first decade of this century, where transposition measures of the Data Protection Directive resulted in a heterogenous patchwork of national laws allowing Big Data to choose places of establishment in Europe based on stringency of Member State laws and especially national supervisory bodies. Arguably, the decision to go with a regulation instead of directive was chiefly motivated by case law appearing before the ECJ and the ECHR revealing unacceptable variance in interpretation of EU data protection standards in the national legal systems of Member States.

What the GDPR and the NIS teach for future AI legislative efforts can be divided into three main areas. The first one is identifying the main actors – the developers and operators of AI technology. The future European AI regulation needs to define criteria for AI, whatever it may be – number of potential users, risk to the rights and freedoms of those users, impact on the market or economy of a Member State or the whole Union - to distinguish operators of AI technology from other actors in the information society industries. At this point, it seems that a more prudent course would be to choose the NIS model for identifying essential service operators, rather

than go with an elegant but uncompromising definition such as those contained in the GDPR. The later would attract criticism from business and investors and require adopting additional measures to limit application to cases where such definitions result in stifling and onerous obligations for the developing market. An example of such provisions, although very limited in practical application, can be found in Article 30 of the GDPR for micro and small business. Similar but in effect broader provisions might either undermine the whole system and/or create incentives for large companies to outsource AI activities to start-ups in order to circumvent and evade the regulation altogether.

The second area would be the principles of AI processing, which should be established analogously to the basic principles of personal data processing, including accountability. Accountability as defined by Article 5 of the GDPR as well as the Article 4 of the Regulation 2018/1725 is a key concept and a basic principle of data protection.

The principle of accountability, building on the six other basic principles (lawfulness, fairness and transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality) ensures that all actors that gain access or hold personal data are liable for any personal data breaches including disclosure, damage, loss or unavailability of personal data entrusted to them on any lawful basis. The General Data Protection Regulation explicitly regulates the data subject right to damages as well as punitive administrative fines on the scale measured not just by tens of millions of euros, but also in 2 or 4 percent of the annual worldwide turnover, whichever is higher, as regulated by the Article 83 of the GDPR.

The High-Level Expert Group on AI Ethics Guidelines for Trustworthy Artificial Intelligence explicitly number seven key requirements for AI systems to meet to be deemed trustworthy. Several of these requirements, such as demand for human agency and oversight, transparency, non-discrimination and fairness and accountability could be easily adopted as principles of trustworthy AI processing.

In turn these principles may reflect in data subject rights in a manner similar to how personal data processing principles now explicitly established by the GDPR reflect in data subject rights in Articles 12 to 22 (the rights to be informed, to have access to personal data, rectification and erasure, restriction of processing, data portability, right to object and automated individual decision making).

The third area would define the required the technical and organisational protection measures to ensure safe and secure processing such as those regulated by Articles 24 through 39 of the GDPR or the Articles 14 through 20 of the NIS Directive. The discussion on exact compliance mechanisms that would safeguard the protection of individual's rights as well as ensure processing in line with envisaged AI processing principles would significantly exceed the confines of this paper, however, certain

good practices have become apparent in the course of application of the GDPR and the national laws in the area of information security regulation.

Independent supervisory bodies are vastly more efficient and useful in both conducting oversight and helping individuals secure their rights and constructively recognizing the best practices and preparing guidelines and opinions how to apply the regulation in practice.

Another key component is ensuring that adequately educated and knowledgeable professionals fill out roles such as data protection officers (GDPR, Articles 36-39) or information security advisers, as regulated by the Article 25 of the Croatian Information Security Act. Interdisciplinary understanding of data protection law and practice and information security standards is obligatory for individuals filling out these roles in order to be able to perform tasks mandated by the Article 39 of the GDPR or, in the Croatian case, Articles 25 and 26 of the Information Security Act.

At the same time, according to market research, the EU Member States potentially lack hundreds of thousands of information security professionals (Ashford, 2017). While this current lack of experts may negatively influence application of legislation currently in force and discourage legislators from adopting similar measures in future AI regulation, it may also have a positive effect in stimulating the academia and education business to develop new interdisciplinary programs and life-long learning courses to train the next generation of information security experts and help existing workforce transition into new jobs facilitated by the information society economy.

Ethical Principles in Trustworthy AI

The second component is based on the ethical principles that should ensure development of an ethical AI. These principles are examined by a sub-field of applied ethics – AI ethics – which focuses on ethical issues that arise with development, deployment and use of Artificial Intelligence. The Ethical Guidelines (European Commission, 2019) establish four ethical principles, based on the recognized fundamental rights, that future AI systems need to adhere to in order to be recognized as trustworthy. These principles, as stated by the Guidelines, are:

- Respect for human autonomy
- Prevention of harm
- Fairness
- Explicability

From the regulatory perspective, in order to analyse the impact of these principles on future regulation of artificial intelligence and based on the lessons learned in the

process of developing data protection and information security regulations, the ethical principles require a closer look and analysis of connection to established regulation principles.

Principle of Respect for Human Autonomy

The Guidelines establish the principle of human autonomy as a reflection of fundamental rights which enshrine the respect for individual's freedom, self-determination and the ability to participate in democratic processes.

Even though most AI systems in use are merely specific expert systems applied to a limited problem, the AI technology in use today already has ability to augment and complement human intellectual skills.

These technologies may become compromised or subverted in order to deceive, coerce and otherwise manipulate individuals by carefully selecting information, applying advanced digital censorships methods and using communication technologies to influence their understanding of the world. Even in this early stage of machine learning, individuals have already experienced some of these effects while using social networks and other Web 2.0 services.

In order to ensure this principle, the Guidelines recommend human-centric design principles and opportunity for individuals to express their choices and securing human oversight over the data processes done by AI systems. This immediately draws comparison with established data subject rights, such as the right to object to processing, profiling or automated individual decision making as regulated by Articles 18, 21 and 22 of the GDPR. The problem here is that using AI indeed in its essence means using automated decision making.

Principle of Prevention of Harm

AI systems should operate in a way that will not harm or adversely affect individuals. The Guidelines stipulate that AI data processing needs to ensure dignity as well as physical and mental integrity of individuals and should operate in a safe and secure manner. The examples of asymmetrical position of power the Guidelines use to illustrate the need to ensure secure and safe processing, such as relationships between employers and employees, the government and citizens or the position of the consumer, illustrate the need to ensure secure and safe processing in a way that is reminiscent of the provisions of data protection law.

The Guidelines also explain that prevention of harm does not apply only to individuals, but also to the natural environment and all living beings. Translating this

principle into a sustainable and efficient legal norm will probably represent the greatest law-making challenge out of the principles described herein. Harm or damage to individuals, let alone society or natural environment in general may come as a result of factors which may be very difficult to predict. Too restrictive interpretation of this principle on the other hand will significantly if not totally prohibit AI development.

Principle of Fairness

According to the Guidelines, the focus of the fairness principle is preventing unfair bias, discrimination or stigmatization of individuals and groups. Again, this line of thinking is reminiscent of the discussions and legal solutions established in the legal framework of data protection. For example, considering the right to object to profiling and automatic decision making in Article 22 of the General Data Protection Regulation, the Regulation contains a balancing provision that checks for necessity of such processing if it produces legal effects concerning the data subject that may be useful in considering future AI regulation.

The Guidelines expand on this notion and imply that the principle of fairness in developing and deployment of AI systems should ensure values such as equal and just distribution of benefits and costs, prevent unfair biases and facilitate equal opportunity in accessing education, goods, services and technology and prevent deception or impairment of individuals concerning their choices.

Data subjects need to have a practical ability to contest and seek protection and redress against AI decisions. The Guidelines again underline the need to include humans operating AI technology in the decision making and control loop, the need to clearly identify the decision making entity and sufficiently explain how the AI system reaches decisions.

Principle of Explicability

Explicability is crucial for establishing and preserving user trust regarding the function of artificial intelligence systems, as understanding of the capabilities, purpose and decision-making process need to be clear to the individuals affected by their actions.

The problem, which the Guidelines acknowledge, is that current machine learning algorithms often produce results without the possibility to backtrack the path taken by the software agent to arrive at the result. The Guidelines suggest using other explicability measures such as traceability, auditability of the underlying information system as well as transparent communication of system capabilities, however,

these measures may not be enough to adequately ensure the required degree of transparency, which will be dependent on the actual context of the processing and the potential impact and severity of the results if they are erroneous or otherwise inaccurate.

From the wording of these principles and the interpretations offered by the Guidelines it is apparent that some of these principles in certain situations may come into mutual conflict, unsurprising as they reflect the fundamental rights of individuals in the European Union which themselves have an established history of balancing in their everyday application and judicature of national and European courts of law.

Regulation of Technical Robustness through Risk Management and Information Security

Finally, the Guidelines stipulate the third component – the need to ensure the development of a robust artificial intelligence systems which will perform their processing in a safe, secure and reliable manner with considerable resilience to adverse impacts, notably the rise of information security incidents and cybercrime in general.

The European Network and Information Security Agency (ENISA) and EUROPOL have for the past decade regularly published yearly research into the trends of information security incidents and detected attacks against information systems, applications and data (ENISA 2018). The research shows an increase in low-skilled, massively distributed attacks committed increasingly by perpetrators without advanced knowledge and understanding of information systems. These attacks continue against a vast majority of European enterprises and business organisations, rising at rate up to 40% year on year (Juncker, 2017).

Technical robustness of future AI developed in Europe requires that these systems be developed with risk minimisation in mind. The Guidelines suggest this may be accomplished with development that ensures resilience to attack and adequate security, planning for service fallback and safety, and accuracy, reliability and reproducibility of results of AI processing.

Risk management plays an important role in recent regulation. Both the General Data Protection Regulation and the NIS Directive feature distinct mechanisms relying on risk management as a recognized information security practice, in fact, information security is a fundamental principle on which the new European system of personal data protection is now explicitly based. The GDPR references the term information security only once, in the Recital 49 when deliberating on the appropriate legal basis for processing of personal data by public authorities, computer emergency response teams (CERTs), computer security incident response teams

(CSIRTs), providers of electronic communications networks and services and by providers of security technologies and services (GDPR, 2016).

However, there are many provisions in the GDPR that indicate the importance of information security for personal data processing. The Regulation references the concept of risk and risk management over seventy times. One of the key data protection principles, the principle of confidentiality and integrity of personal data processing, is basically a modification of the standard CIA triad of information security – Confidentiality, Integrity, Availability (Nissenbaum, 2005:61) The obligations of data controllers in incident response mirror the obligation of critical service providers etc. The purpose of this application of information security concepts is to ensure the adequate level of accountability of data controllers and processors.

Additionally, the principle of accountability now demands that data controllers need to document and demonstrate accountability. This principle is elaborated and developed in numerous GDPR provisions, especially concerning the obligations of the data controllers and processors, such as those regulated by Articles 24, 25, 28, 30, 32, 33 and 34 of the GDPR.

The Article 24 lays out responsibilities of the data controller, including implementation of technical protection measures implementation of appropriate data protection policies and possible demonstration of compliance through certification mechanisms and approved codes of conduct. The technical measures may range from simple solutions to prevent unauthorized access and distribution of data from relevant information systems to complex data protection management solutions that integrate data discovery, endpoint security and other functionalities available in the information security solutions market. Additionally, the organizations processing personal data are required to establish procedures when handling personal data and mechanisms to supervise and document their behaviour to prove their accountability.

Article 25 regulates another novel concept in the GDPR – *data protection by design and by default.* The provisions of the Article demand that when deciding on the appropriate safety measures the data controller needs to take into account the developments in information technology, security and business practices. The controller also needs to assess the purpose and scope of processing, cost of implementing appropriate measures and potential risks – how likely the potential data breaches are and what would be the severity of their impact on rights and freedoms of individuals.

The provisions of Article 25 offer a balanced approach that weighing both the dangers and the costs. On the base of that assessment the data controller has to implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the pro-

cessing. In other words, from the start the controller who decides on the purpose and means of processing needs to choose adequate measures to ensure secure and safe processing and minimize risks for the data subjects.

Article 28 elaborates on the work of the WP29 group and national supervisory bodies in regulating the transfer of personal data to third countries resulting in the creation of standard contractual clauses. These experiences have helped frame the Regulation obligations for data protection contracts between data controllers and data processors. The article stipulates the obligatory provisions and clauses of such contracts, emphasizing the duties of the controller to ensure safe and secure processing and setting the standards of accountability in choosing a processor complying with appropriate safeguards. The controllers need to ensure that their processors and sub-processors are contractually obliged to comply with Regulation standards.

As the number of information security incidents continues to rise, and the impacts of data breaches targets a rising number of data subjects both in volume as in severity, the data controllers need to prepare for potential incidents. This includes both the technical and organisational means to discover the incidents as well as to ascertain risk and damage, mitigate threats and recover data and finally resume normal processing operations and report to supervisory bodies and data subjects if need be. The Regulation lays out these obligations in Articles 32-34.

These provisions, especially obligations regarding documenting the processing activities and implemented technical protection measures of the data controller impose controls over data processing activities. Controllers and processors are required to document their handling of personal data and implementation of measures to ensure safe and secure processing, export of data to other organizations and record potential incidents or data breaches.

These obligations of data controllers in case of a data breach towards the supervisory bodies and the data subjects whose data has suffered a breach are the cornerstone of the system of accountability created by the GDPR establishing a system of almost objective, strict liability (Van Alsenoy, 2017). Very similar provisions are present in the NIS Directive and its national transpositions measures.

This approach has received a substantial amount of criticism, usually from the position of global multinational corporations and foreign powers without a comparable national data protection legal framework. Typically, the GDPR requirements are found to be oppressive, anti-business, intimidating and dangerous to established business models, usually without considering the benefits for rights and freedoms of data subjects. The same narrative can already be discerned in the popular media, if not already in comparative legal literature, about the intentions of the EU regulators to develop a legal framework for AI.

Conclusion – What Lessons for AI Regulatory Framework

From the publicly available documents, it seems the EU Commission has recognized that preserving trust of its citizens and individuals residing in the EU is the single most important task of any future AI regulation in Europe. Having the practical ability to guard fundamental rights in light of the emerging artificial intelligence technologies is vital for future legal, economic and political sovereignty of the European Union.

Additional goals of such regulation would be ensuring the legal framework is not prohibitively constrictive as to prevent research and development of AI technologies and new, innovative products and services for the European and global market. The AI regulation, like GDPR, should create opportunities for new jobs, education and professional specialization to adapt the workforce to the needs of information society.

Recognizing the positive effect recently introduced regulation such as the GDPR, the NIS Directive and other regulation that seeks to balance the established level of fundamental rights with the needs of the information society has had on comparative legislative practice, the European lawmakers would do good to adopt successful ideas and solutions from these sources. These laws have brought forth into the EU legal system and the legal systems of the Member States concepts well established in the judicature of the European Court of Human Rights and the Court of Justice of the European Union. They have also introduced mechanisms proven by decades of practice in information security industry and these practices would serve well as an operating premise for the AI regulation.

The first step in this regard has already been taken – distillation of ethical principles into principles of AI design in the way reminiscent of personal data protection principles developed in EU during the last quarter of the century and now explicitly recognized by the General Data Protection Regulation. The second will be more political in nature – choosing between a regulation or a directive.

Additionally, the AI legal framework needs to strengthen compliance mechanisms and establish competent independent supervisory regime along side with promotion of new, specialized experts to tackle the operational, practical issues of oversight over developers and providers of AI technology and services.

REFERENCES

Berendt, B., Littlejohn, A. & M. Blakemore (2020) "AI in education: learner choice and fundamental rights", *Learning, Media and Technology*, 45 (3), 312–324. doi: 10.1080/17439884.2020.1786399.

Butarelli, G. (2016) "The EU GDPR as a clarion call for a new global digital gold standard", *International Data Privacy Law,* 6 (2), 77–78. https://doi.org/10.1093/idpl/ipw006.

Chen, L., Qiao, Z., Wang, M., Wang, C., Du, R. & H. E. Stanley (2018) "Which Artificial Intelligence Algorithm Better Predicts the Chinese Stock Market?", *IEEE Access*, 6, 48625–48633. doi: 10.1109/ACCESS.2018.2859809.

Dilsizian, S. E. & E. L. Siegel (2014) "Artificial Intelligence in Medicine and Cardiac Imaging: Harnessing Big Data and Advanced Computing to Provide Personalized Medical Diagnosis and Treatment", *Curr Cardiol Rep*, 16, 441. doi: 10.1007/s11886-013-0441-8.

Grisham, J., Samtani, S., Patton, M. & H. Chen (2017) "Identifying mobile malware and key threat actors in online hacker forums for proactive cyber threat intelligence, *[IEEE 2017 IEEE International Conference on Intelligence and Security Informatics (ISI) – Beijing*. China.

Katulić, T. (2015) "Protection of Computer Programs in Comparative Law: Current Issues and Development Perspective", *Zbornik Pravnog fakulteta u Zagrebu*, 65 (29), 237–262.

Koščík, M., Harašta, J., Kyncl, L., Myška, M., Polčák, R. & V. Stupka (2017) *European ICT Law 2017: Casebook of ICT Law*. Brno: Masaryk University.

Lessig, L. (1999) "The Law of the Horse: What Cyberlaw Might Teach", *Harvard Law Review*, 113 (2), 501–549.

Malgieri, G. & B. Custers (2018) "Pricing privacy – the right to know the value of your personal data", *Computer Law & Security Review,* 289–303.

Moore, G. E. (1965) "Cramming More Components onto Integrated Circuits", *Electronics*. New York: McGraw-Hill.

Nissenbaum, H. (2005) "Where Computer Security Meets National Security", *Ethics and Information Technology,* 61–73. doi: 10.1007/s10676-005-4582-3.

Ontañón, S., Synnaeve, G., Uriarte, A., Richoux, F., Churchill, D. & M. Preuss (2013) "A Survey of Real-Time Strategy Game AI Research and Competition in StarCraft", *IEEE Transactions on Computational Intelligence and AI in Games*, 5 (4), 293–311. doi: 10.1109/TCIAIG.2013.2286295.

Prins, C. (2016) "When Personal Data, Behavior and Virtual Identities Become a Commodity: Would a Property Rights Approach Matter", *SCRIPT-ed* 3 (4), 270–303.

Safari, B. (2016) "Intangible Privacy Rights: How Europe's GDPR Will Set a New Global Standard for Personal Data Protection", *Seton Hall Law Review*, 47 (3), 6. https://scholarship.shu.edu/shlr/vol47/iss3/6.

Sebag, M. (2014) "A Tour of Machine Learning: An AI Perspective", *AI Communications*, 27 (1), 11–23.

Solomonoff, R. J. (1985) "The Time Scale of Artificial Intelligence:Reflections on Social Effects", *Human Systems Management*, 149–153.

Taddeo, M. & L. Floridi (2018) "Regulate artificial intelligence to avert cyber arms race", *Nature*, 556, 296–298. doi: 10.1038/d41586-018-04602-6.

Terry, N. (2019) "Of Regulating Healthcare AI and Robots", *Yale Journal of Health Policy, Law and Ethics, Yale Journal of Law and Technology*. doi: 10.2139/ssrn.3321379.

Van Alsenoy, B. (2017) "Liability under EU Data Protection Law: From Directive 95/46 to the General Data Protection Regulation", *Journal of Intellectual Property, Information Technology and E-Commerce Law*, 271.

Vochozka, M., Klieštik, T., Klieštiková, J. & G. Sion (2018) "Participating In A Highly Automated Society: How Artificial Intelligence Disrupts The Job Market", *Economics, Management, and Financial Markets*, 13 (4), 57–62.

Websites

Bassot, E. & W. Hiller (2019) "The Juncker Commission's Ten Priorities: An End of Term Assessment", *European Parliament*. https://www.europarl.europa.eu/RegData/etudes/IDAN/2019/637943/EPRS_IDA(2019)637943_EN.pdf.

Columbus, L. (2019) "Top 25 AI Startups Who Raised the Most Money in 2019", *Forbes.com*. https://www.forbes.com/sites/louiscolumbus/2019/12/22/top-25-ai-startups-who-raised-the-most-money-in-2019.

ENISA Threat Landscape Report (2018) *European Network and Information Security Agency*. https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2018.

European Commission (2012) *Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)*. https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52012PC0011.

European Commission (2016) *Communication from the Commission to the European Parliament, the Council, the European economic and social committee and the Committee of the regions - Online Platforms and the Digital Single Market Opportunities and Challenges for Europe*, COM/2016/0288. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016DC0288&from=EN.

European Commission (2017) *State of the Union 2017 – Cybersecurity: Commission scales up EU's response to cyber attacks*. http://europa.eu/rapid/press-release_IP-17-3193_en.htm.

European Commission (2018) *European Parliament Directive on representative actions for the protection of the collective interests of consumers, and repealing Directive 2009/22/EC*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018PC0184.

European Commission (2019) *Ethics Guidelines for Trustworthy AI*. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

European Commission (2019) *Policy and investment recommendations for trustworthy Artificial Intelligence*. https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

European Commission (2020) *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust*. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.

European Communities (1995) "Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data (Data Protection Directive)", *Official Journal of the European Union*, L 281, 31–50. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31995L0046.

European Communities (2000) "Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on Electronic Commerce)", *Official Journal of the European Union*, L 178, 1–16. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32000L0031.

European Union (2012) "Charter of Fundamental Rights of the European Union", *Official Journal of the European Union*, *C 326, 391–407. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=OJ%3AC%3A2012%3A326%3ATOC.*

European Union (2016) "Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union, (Network and Information Security Directive)", *Official Journal of the European Union*, L 194, 1–30. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016L1148.

European Union (2016) "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 on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)", *Official Journal of the European Union*, *L 119, 1–88. https://eur-lex.europa.eu/eli/reg/2016/679/oj*.

European Union (2108) "Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC", *Official Journal of the European Union*, L 295/39. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R1725.

The Croatian Parliament (2007) "Information Security Act", *Official Gazzette of the Republic of Croatia* 79/07. https://www.uvns.hr/UserDocsImages/en/dokumenti/info-security/Information-Security-Act.pdf.

The Ministry of Interior (2018) "Act on the implementation of General Data Protection Regulation", *Official Gazzette of the Republic of Croatia* 42/18. https://mup.gov.hr/personal-data-protection/124.

Von der Leyen, U. (2019) "A Union that strives for more: My agenda for Europe, Political Guidelines for the Next European Commission 2019-2024", *EU Commission.* https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf.

Medij. istraž. (god. 26, br. 2) 2020. (9-28)

Prema regulaciji umjetne inteligencije (UI): uvidi i iskustva iz regulacije zaštite osobnih podataka i informacijske sigurnosti

Tihomir Katulić

SAŽETAK

Razvoj naprednih informacijskih tehnologija omogućio je da nakon desetljeća teorijskih razmatranja u praktičnu primjenu uđu prvi oblici strojnog učenja kao koraka prema razvoju umjetne inteligencije. Tranzicija u postindustrijsko, informacijsko društvo otkrila je važnost podataka kao važnog resursa na čijoj se obradi temelje nove inovativne informacijske usluge. Europski je zakonodavac u prethodnom razdoblju usvojio niz važnih zakona kojima je cilj zaštititi prava pojedinca i regulirati obveze davatelja takvih usluga kako bi se osigurala sigurna obrada podataka.

Povelja o temeljnim pravima Europske unije, jedan od temelja suvremenoga europskog sustava ljudskih prava, sadrži značajne kontrole i ograničenja koja će utjecati na razvoj i svrhu buduće regulacije umjetne inteligencije na području Europske unije. Budući propisi trebat će sadržavati i pridržavati se usvojenih europskih pravnih standarda oko zaštite temeljnih prava pojedinaca u Uniji. Etičke smjernice Europske komisije predstavljaju korak prema usvajanju etičkih principa, temeljenih na prepoznatim temeljnim pravima, koji će biti obvezujući za informacijske sustave zasnovane na umjetnoj inteligenciji. Cilj je ovog rada istražiti i analizirati rješenja postojećih europskih propisa iz područja zaštite osobnih podataka i informacijske sigurnosti kao i dosad objavljenih dokumenata Europske unije o budućoj regulaciji umjetne inteligencije te ponuditi rješenja de lege ferenda. Rad se zasniva na komparativnom prikazu i analizi odabranih izvora i odredbi europskog i nacionalnog zakonodavstva, pravne književnosti i drugih znanstvenih izvora.

Ključne riječi: umjetna inteligencija, zaštita osobnih podataka, informacijska sigurnost, informacijsko društvo