Abstract

Transparency is a central and significant concept in law and important when creating trust. Therefore, it is always challenging to find the right balance between established principles and a new technology that undergoes complicated development processes.

The main concern when it comes to AI technology is the difficulty for individuals to challenge automated decisions, especially when the decision-making process and the data used as basis for the algorithms involved are not clearly stated and publicly available.

There was hope the EU led proposed new AI Act would contain clear regulations regarding transparency obligations making AI systems more trustworthy for impacted individuals but this was not the case. Instead, transparency issues have been left to self-regulation by developers creating a potential power imbalance between private companies developing AI technology and public authorities deploying it. The AI Act's chosen risk-based approach is not future proof as it does not set clear criteria in regards to the classification of future AI technology creating additional difficulties.

Even though the AI Act is a good starting point with the potential to become a globally accepted regulatory framework, it fails to address the fundamental power imbalance between developer and deployer of AI systems and those affected by it, making it a regulation for companies, not for people.

European Judicial Training Network (EJTN)

THEMIS Competition 2022

Semi-final D: Judicial Ethics and Professional Conduct - TH/2022/04

The Concept of Transparency in the Proposed European Artificial Intelligence Act[1] and the Consequences for Governmental Institutions, Government Actions and Public Services

by

Team Germany:

Jessica Kiel, LL.M.-/Ph.D.-Candidate

Jessica Reisinger, LL.M. *(Budapest)*

Julia Werner, Ph.D., LL.M. *(Waikato)*

Tutor: Valerie Datzer, Maître en droit (*Bordeaux*)

[1] European Commission *Proposal for an Artificial Intelligence Act* (AIA) of 21 April 2021, COM(2021) 206 final. In the following: AI Act.

1. Introduction

Technology has improved the efficiency and accuracy of public and private legal services over the last decades with more and more processes being 'automated' in the day-to-day life of legal professionals in all areas of the law. And now Artificial Intelligence (AI), designed to reduce the burden on humans, optimize processes, complete tasks faster and more accurately, while minimizing sources of human error, promises to change the whole economy and the law respectively.

Over the last couple of years, AI has raised very extensive and profound questions of value and human nature triggering a rush to draw up codes of ethics for AI together with technical standards concerning ethics and safety. The most concerning question remains, whether AI might surpass human control and comprehension and what will be the best preparation for such a scenario.[2] This triggered the recent development of many principles and guidelines for legal and ethical AI usage by different organisations and governments around the world, including the OECD's *Principles on Artificial Intelligence*[3], the *Universal Guidelines for AI*[4], the *European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*[5] and the most recent EU Commission Proposal for an *Artificial Intelligence Act* (AI Act)[6]. But because the use of AI systems and their impact does not stop at national borders, global cooperation and coordination of AI regulations is crucial to avoid a focus on global competitiveness.[7]

The AI Act has to consider the concept of transparency and at the same time ensure fundamental rights and legal ethics when finding a balance between social interests in innovation and more efficient delivery of public services and governmental actions.

---

[2] Boddington, Paula 'Normative Modes: Codes and Standards', in Markus D. Dubber, Frank Pasquale, and Sunit Das (eds.), *The Oxford Handbook of Ethics of AI* (2020) 125, at 126.
[3] OECD, *Recommendation of the Council on Artificial Intelligence* **(**adopted 22.05.2019), available at https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.
[4] The Public Voice Org, *Universal Guidelines for Artificial Intelligence*, (2018), available at https://thepublicvoice.org/ai-universal-guidelines/.
[5] European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment*, available at https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.
[6] European Commission COM(2021) 206 final 2021/0106(COD).
[7] Ulnicane, Inga *et al.*, 'Good governance as a response to discontents? Déjá vu, or lesson for AI from other emerging technologies' 46 *Interdisciplinary Science Reviews* (2021), 71, at 81.

This paper provides a brief overview of the ethical and legal challenges the use of AI technology brings for governmental agencies in different areas of the law with respect to regulations proposed in the AI Act. It describes how AI technology is affected by data acquisition and algorithm transparency and analyses the consequences for Administrative Law practices, especially in regards to the rights of citizens, and Criminal Law practice focusing on the example of facial recognition. It concludes that not making clear statements regarding transparency and leaving it to self-regulation ultimately leads to a power imbalance between developer and deployer of AI technology.

2. The Concept of Transparency and the Effects of Human Judgment on Algorithms

AI is a wide field and as a discipline within computer science, depending on the usage it can be many things and has been described as:

> 'cross-disciplinary approach to understanding, modelling and replicating intelligence and cognitive processes by invoking various computational, mathematical, logical, mechanical, and even biological principles and devices […] [i]t forms a critical branch of cognitive science since it is often devoted to developing models that explain various dimensions of human and animal cognition.'[8]

The most recent attempt on a legal definition was set out by the European Commission in the AI Act defining AI systems in Article 3 (1a) and Annex I as meaning

> […] software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with'
> […] (a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;

[8] Frankish, Keith and Ramsey, William M. (eds.) *The Cambridge Handbook of Artificial Intelligence* (2014), at 1.

(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

(c) Statistical approaches, Bayesian estimation, search and optimization methods.

This rather broad definition allows for flexibility in view of the rapid technical progress in AI systems, but it has been criticised due to potential legal uncertainty for developers, operators, and users of AI systems.[9] One reason for the criticism is that the definition does not acknowledge that AI technology is an intangible product or more likely a service which learns and changes along the way depending on the algorithm it was based upon and the use it was created for.[10]

A. Human decision making and algorithmic transparency

Algorithms are the 'heart' of any AI system and human judgment is unavoidably exercised in designing such systems and programming the algorithm software.[11] Algorithms are understood as being 'strictly rational concerns, marrying the certainties of mathematics with the objectives of technology'.[12] Algorithms embody a wide range of judgment at the level of design '[i]n reality […] a great deal of expertise, judgment, choice, and constraints are exercised in producing algorithms.'[13] The big seven judgments that have to be made when creating AI decision making systems are described by *Kitchin* as follows:

1. How to characterise the relevant task;
2. Translation of the task or problem into a structured formula with an appropriate rule set;

[9] Bomhard, D. and Merkle, M., 'Europäische KI-Verordnung. Der aktuelle Kommissionsentwurf und praktische Auswirkungen', 6 *Recht Digit.* (2021), 276, at 278.
[10] Edwards, L., *Regulating AI in Europe: four problems and four solutions* (2022), available from Ada Lovelace Institute https://www.adalovelaceinstitute.org/report/regulatingai-in-europe/.
[11] Spaulding, Norman W., 'Is Human Judgment Necessary?: Artificial Intelligence, Algorithmic Governance, and the Law' in Markus D. Dubber, Frank Pasquale, and Sunit Das (eds.), *The Oxford Handbook of Ethics of AI* (2020).
[12] Kitchin, Rob, 'Thinking Critically About and Researching Algorithms' 20 *Information, Communication and Society* (2017), at 17.
[13] Kitchin, at 18.

3. Translating this pseudo-code into a source code that when compiled will perform the task or solve the problem;
4. how to deal with time and resource constraints for the design and execution of the project;
5. the choice and quality of training data;
6. how to deal with 'requirements relating to standards, protocols and the law'; and
7. how to manage 'conditionalities related to hardware, platforms, bandwidth, and languages.'[14]

This shows that when breaking down AI technology simply into the 'if x / than y' Boolean logic[15] and algorithm syntax[16], it requires a series of very complex decision making to create such systems. Typically, software engineers will set thresholds choosing the value that seems reasonable but users might not have knowledge of those thresholds or might not even be aware that such thresholds exist. [17] For value and respectively ethical judgments, there would be a certain elevated need to transparently communicate such thresholds as they are the starting point when it comes to explainability and the understanding of how AI automated-decisions have been made. The choices made when designing AI systems is reflected in its decision-making but what is known so far suggests that many such systems process information differently from what humans would usually do, even though the data on hand is the same.[18]

This adds to the problems of making algorithms truly transparent and decisions made with or by AI technology explainable along the line as the technology learns from the data it is given depending on the algorithm and the use it was designed for. It also adds to transparency issues that have to be addressed when discussing regulations on AI technology, especially when it comes to the right amount of transparency in respect to the different stakeholders.

B. General Transparency Issues with AI technology and ADM systems

[14] Kitchin, at 17.
[15] The idea that all values are either true or false. For example, if certain conditions are met, the value of a specific query is true or vice versa.
[16] Syntax in computer programming means the rules that control the structure of the symbols, punctuation, and words of a programming language.
[17] Spaulding, at 6.
[18] Selbst, Andrew D. and Barocas, Solon, 'The Intuitive Appeal of Explainable Machines' 87 *Fordham Law Review* (2018) 1085, at 1089–1090.

In many areas AI technology and especially automated decision-making (ADM) systems already replace or complement human thought and decision making but it remains crucial to ensure the affected parties can exercise their right of an explanation to why a decision was made because transparency is a key value and a fundamental principle of the *EU Treaties* and *Charter of Fundamental Rights of the European Union*[19]. Transparency is also essential for general public law which '*inter alia* requires public bodies to give reason for their decision involving civil consequences to individuals' but AI systems also need opacity for a wide range of reasons like trade secrets and copyright related issues when it comes to developing such systems and guaranteeing data safety.[20]

While there is no universal definition of the principle of transparency, it can be defined in light of Article 1 of the *Treaty on the European Union* (TEU) [21], as 'principle according to which decisions should be taken as openly, comprehensible and verifiable by the public and the affected persons, including in particular the basis of decision, relevant decision criteria and the overall decision finding process'.

The overall decision finding process as described in the above definition ultimately collides with AI technology when it comes to data acquisition and use as well as the development process of the specific system providing the basis for incorporated algorithms. If those algorithms are flawed or the data-sets it was trained on are erroneous, violation of fundamental rights could be the result making it next to impossible to verify the decision criteria and the decision finding process if there are no transparency standards or regulations regarding the system used. [22]

[19] European Union Charter of Fundamental Rights of the European Union. Official Journal of the European Union C83. Vol. 53.
[20] Wynne, B. 'Risk and Environment as Legitimatory Discourses of Technology: Reflexivity Inside out?' 50 *3 Current Sociology* (2020), 459–477; Alì, G. S., and Yu, R. 'Artificial Intelligence Between Transparency and Secrecy: From the EC Whitepaper to the AIA and Beyond' 12 *3 European Journal of Law and Technology* (2021), 1.
[21] Consolidated version of the Treaty on the European Union [2012] OJ C326/13.
[22] Varošanec, Ida (2022): 'On the path to the future: mapping the notion of transparency in the EU regulatory framework for AI' International Review of Law, Computers & Technology (2020) published online available at https://doi.org/10.1080/13600869.2022.2060471 at 3.

But at the same time, humans may also fall short of explanation for their decision and they might be unaware of all the factors that led to their final decision depending on experiences and other reasons. The question is how much perfection can be expected of an AI system and what are the transparency requirements in regards to different systems and stakeholders.

C. Specific Transparency Requirements

What each group needs to know in regards to transparency is depending on their rights and obligations regulators require a short non-technical description of the algorithmic tool and the reason of its use together with an explanation how the tool works and the data it uses (threshold, value judgments) to minimise the risk of erroneous decisions.[23] This is especially important for public bodies using AI technology as they need to be able to observe principles of good administration and government as well as the duty to give reasons in their decision making. Therefore, they require information on what elements were considered reaching a certain decision (threshold and value judgments made during the designing process of the AI tool) to explain the reasoning to the affected person.[24] Transparency for consumers might require access to testing data and results, to ensure the concerned product fitted with an AI tool is safe to use.[25]

In regards to AI technology and ADM systems, transparency requirements should focus on making the knowledge available that is necessary for each group to understand the decision-making process the AI system was designed for, especially the decisions surrounding the used data and thresholds and value judgments but not so much details of the technical development process. Full transparency of the latter would enable 'those who know enough about the technology [to] obtain goods or services unfairly' or even enable them to corrupt or misuse data.[26] Therefore, explainability is needed under transparency standards for each group of stakeholders to ensure the duty of giving reason and minimise the risk of unexplainable

[23] Kingsman, N. *et al.*, 'Public Sector AI Transparency Standard' (2021) available at SSRN 3986213 http://dx.doi.org/10.2139/ssrn.3986213.
[24] Opdebeek, I., and De Somer, S. 'The Duty to Give Reasons in the European Legal Area: A Mechanism for Transparent and Accountable Administrative Decision-Making? A Comparison of Belgian, Dutch, French and EU Administrative Law' 2 Rocznik Administracji Publicznej (2016), 97, at 96.
[25] Weller, A. 'Challenges for Transparency' (2019) available at https://arxiv.org/pdf/1708.01870.
[26] Diakopoulos, N. 'Accountability in Algorithmic Decision Making' 59 *2 Communications of the ACM* (2016), 56, at 56–62.

decisions. Under the standards of transparency, it is most important to enable involved and affected parties to challenge decisions, particularly in light of the data and algorithms used as basis for the decision-making process.

A lack of transparency and especially algorithmic and data transparency would make it difficult for involved parties to identify errors, to contest and potentially demand correction of, and to ultimately receive compensation for erroneous decisions.[27] This would have the potential to affect fundamental rights like fair trial and due process, effective remedies, social rights and access to public services as well as rights to free elections. The most vulnerable groups likely to be affected are people that have been denied jobs, refused loans or other public and private services due to an algorithm-based system they fall through and cannot contest due unexplainable decision-making processes.[28]

Consequently, it is of the utmost importance that AI technology and ADM systems come under an equivalent scrutiny as conventional decision making by humans to ensure transparent decision making and explainability of such. Especially when it comes to erroneous decisions, the explanation of the reasoning needs to be easily accessible, explainable and understandable for all involved parties.

The challenge that arises is to make AI technology more trustworthy by finding the right balance between industrial and transparency standards without creating uncertainty that would ultimately hamper new innovations. The key to trust is transparency, not only for affected parties, but especially for regulators, policy- and lawmakers, as it is them deciding where and under which conditions AI technology and ADM systems should be used. Especially when used in judicial settings, decisions considered legitimate must be explainable to guarantee their appealability due to the duty to give reason, requiring governmental institutions and other public bodies to make transparent and comprehensive decisions.[29]

3. Requirements for the Principle of Transparency under the AI Act

[27] Varošanec, at 3.
[28] Rodrigues, Rowana 'Legal and human rights issues of AI: Gaps, challenges and vulnerabilities' *Journal of Responsible Technology* (2020) 4, at 8.
[29] Završnik, A. 'Criminal justice, artificial intelligence systems, and human rights' 20 *ERA Forum* (2020), 567–583, available at https://doi.org/10.1007/s12027-020-00602-0.

The EU institutions have progressively concretized the EU's AI agenda through a series of strategic documents. The Commission's work on AI has been reflected in the AI Strategy for Europe 2018[30], the 2020 White Paper on AI[31], and a recently updated Coordinated Plan for AI[32]. And ultimately in April 2021, the Commission presented the legislative proposal for an AI law, following a risk-based approach. The AI Act follows a horizontal approach and, as a measure for the approximation of the laws, regulations and administrative provisions of the Member States, is based on Article 114 TFEU in conjunction with Article 26 TFEU.

This paragraph first explains the division of the AI Act into four risk classes and states in which governmental actions are regularly classified. In the following, it represents abstract terms about how the control of transparency should be designed. Then it analyses which of these abstract provisions have been implemented in the AI Act. Finally, the chapter ends with a brief statement on the AI Act's regulatory framework.

A. The different risk classes in the AI Act

The AI Act is divided into four different risk classes (minimal, limited, high and unacceptable), setting different transparency obligations for providers of AI systems.

1. Minimal and limited risks systems

On the 'minimal risk systems' are minimal transparency obligations imposed[33] (e.g. the free use of applications such as AI-powered video games or spam filters[34]).

[30] Commission Communication of 25.4.2018, Artificial Intelligence for Europe, COM(2018) 237 final,available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN.
[31] Commission White Paper of 19.2.2020 COM(2020) 65 final, available at https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.
[32] European Commission, *Coordinated Plan on Artificial Intelligence 2021 Review* (2020), available at https://digital-strategy.ec.europa.eu/en/policies/plan-ai.
[33] Recital 5.2.2. and 2.3. AI Act.
[34] EU-Kommission, *Künstliche Intelligenz - Exzellenz und Vertrauen* (2022), available at https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_de.

‘Limited risk systems’ (Title IV) are also subject to minimal transparency obligations, ‘for example in terms of the provision of information to flag the use of an AI system when interacting with humans’[35] (e.g. chatbots[36]).

2. *High-risk systems*

The ‘high-risk systems’ are regulated in Article 6 AI Act in conjunction with Annex III.

In Article 6 (1) AI Act an abstract definition of such high-risk systems can be found. An AI system shall be considered high-risk, if both of the following conditions are fulfilled: 'The AI system is intended to be used as a safety component of a product, or is itself a product, covered by the Union harmonisation legislation listed in Annex II’. In addition, ‘the product whose safety component is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment with a view to the placing on the market or putting into service of that product pursuant to the Union harmonisation legislation listed in Annex II.'[37]

Furthermore, ‘AI systems referred to in Annex III shall also be considered high-risk.’[38] In Annex III there is a specific list of high-risk systems: These systems contain critical infrastructure (e.g. transport) where the lives and health of citizens could be put at risk. Education or vocational training, where a person's access to education and professional life could be affected (e.g. assessment of exams). It also includes safety components of products (e.g. an AI application for robotic-assisted surgery); employment, human resource management and access to self-employment (e.g. software to evaluate CVs for recruitment processes); centralised private and public services (e.g. credit scoring, denying loans to citizens); law enforcement that could interfere with people's fundamental rights (e.g. verifying the authenticity of evidence). Furthermore, migration, asylum and border control (e.g. verifying the authenticity of travel documents), administration of justice and democratic processes (e.g. application of legislation to specific facts).[39]

[35] Recital 5.2.2 and 2.3. AI Act.
[36] EU-Kommission, *supra* note 34.
[37] Art. 6 (1) AI Act.
[38] Art. 6 (2) AI Act.
[39] Art. 6 with Annex III AI Act; see also EU-Kommission, *supra* note 34.

3. Unacceptable risk systems

At the top of the 'risk pyramid', the AI Act (Title II) establishes an explicit list of prohibited AI practices which create an 'unacceptable risk'[40] as they violate fundamental rights or go against EU values (e.g. social scoring, 'real-time' remote biometric identification systems). The prohibitions cover AI systems that have a huge potential to 'manipulate persons through subliminal techniques beyond their consciousness or exploit the susceptibilities of specific vulnerable groups in a manner that could cause them or others physical or psychological harm.'[41]

4. Classification of governmental actions

As already mentioned, Article 6 AI Act with Annex III determines when, a risk to the health and safety or fundamental rights of persons is so blatant that the AI system is to be classified as a 'high-risk AI system'. Regularly, in Article 6 (2) AI Act in conjunction with Annex III no. 1, 2, 5-8'[42] are relevant for administrative and criminal law. Therefore, the practice of public authorities regularly falls under high-risk AI systems.[43]

B. Prior and subsequent control of AI systems

The next part should examine how abstract provisions for adequate transparency are implemented in the AI Act. Transparency is essential to understand the decision-making processes of ADM systems and thus to regulate their behaviour.

AI systems can be regulated ex ante and ex post. Through prior transparency checks on AI systems, information on data processing and the functioning of systems is already explored in advance. As a result of ex-post control, explanations and justifications are required from ADM systems, which can then be verified.[44] This reveals how and why a particular decision was made. Because of the benefits of this dual control, a legal framework for AI should include prior and ex-post control.[45]

[40] Art. 5 AI Act.
[41] Recital 5.2.2. AI Act.
[42] Art. 6 (2) in conjunction with Annex III No. 2, 5 and 7 AI Act.
[43] Title III of the AI Act is dedicated to high-risk AI systems.
[44] Zerilli, J. *et al.*, 'Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard?', 32 *4 Philosophy & Technology* (2019), 661–683, at 670 ; Varošanec (2022), 5.
[45] Varošanec, at 5.

1. ex ante-control

According to the EU proposal for an AI Act[46] high-risk AI systems must meet several requirements to be allowed on the market at all. The AI system must include appropriate risk management and systems for risk mitigation.[47] The AI system's data sets must be of high quality[48] in order to minimise overall risks and in particular discriminatory outcomes.[49] Data quality requires that training, validation and testing data are relevant, representative, error-free and complete.[50] Furthermore, the draft stipulates an obligation for detailed technical documentation containing all necessary information about the system and its purpose, so that the public authority can assess compliance with the aforementioned provisions.[51] For instance, a detailed description of the elements of an AI including the design specifications of the system and system architecture need to be included.[52] Moreover, the logging of activities must be ensured in order to be able to trace the results.[53] Furthermore, a high level of robustness, security and accuracy is required.[54]

2. ex post-control

An important aspect for the fulfilment of the transparency obligation after the market launch is human supervision according to Article 14 AI Act. 'Human oversight shall aim at preventing or minimising the risks to health, safety or fundamental rights […]'.[55] Natural persons overseeing the use of AI systems should be able to 'fully understand the capacities and limitations of the system and be able to duly monitor its operation […]'.[56] Additionally the official shall 'remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system ('automation bias') […]'.[57]

In this context, the explicit transparency obligation[58] is also of enormous importance. Pursuant to Article 13 information should be provided for users. Developers will be instructed to 'design

[46] See Title III Chapter 2.
[47] See Art. 9 AI Act.
[48] See Art. 10 AI Act.
[49] See Art. 10 (2) AI Act.
[50] See Art. 10 (3) AI Act.
[51] See Art. 11 AI Act.
[52] See Art. 11 (1) in conjunction with Annex IV No. 2 (b) and AI Act.
[53] See Art. 12 AI Act.
[54] See Art. 15 AI Act.
[55] Art. 14 (2) AI Act.
[56] Art. 14 (4a) AI Act.
[57] Art. 14 (4b) AI Act.
[58] See Art. 13 AI Act.

and develop AI systems in a way that ensures that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately'[59]. According to Article 13 (2) AI Act, AI systems 'shall be accompanied by instructions for use […] that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to users'. Only as a result of transparent operation and the instructions for use provided can a natural person fully understand the capabilities and, in particular, the limitations of an ADM system and thereby also assess its decision.[60] Finally, the Commission, in cooperation with EU Member States, will establish and maintain an EU database of information on high-risk AI systems.[61] Through this database, access to information for citizens - as one aspect of the idea of transparency- should be guaranteed. So the information in the database would be publicly available[62] and the personal data it contains would be limited to what is necessary for the purposes of the database, including the names and contact details of the persons responsible for registering the system.[63] So at the same time, privacy will be ensured.

D. Statement

In summary, the AI Act classifies essential private and public services as well as administration of justice and democratic processes as high-risk AI systems making them subject to several requirements. For instance, high-risk AI systems must ensure some ex ante prerequisites prior to their market authorisation. After the introduction of such systems to the market, they have to fulfill various additional obligations established in the AI Act. The presented requirements[64] explained above are imposed on providers and users.[65] The AI Act was widely criticised for not setting appropriate transparency standards, leaving crucial regulations to AI developers.

[59] Art. 13 (1) AI Act.
[60] See Art. 14 (4a) AI Act.
[61] See Art. 60 (1) AI Act.
[62] See Art. 60 (3) AI Act.
[63] See Art. 60 (4) AI Act.
[64] Title III Chapter 2 AI Act.
[65] See Art. 52 AI Act.

Firstly, it is not clear when data is of high quality. Moreover, it has been criticised that the quality of data cannot ensure a lack of discrimination and bias.[66]
Regarding the obliger (providers and users), developers are instructed to 'design and develop AI systems in a way that ensures that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately'[67] and provide instructions for use that include 'concise, complete, correct and clear information that is relevant, accessible and comprehensible to users'[68]. In this context, it is uncertain what information would guarantee sufficient transparency. It is again unclear, who is to decide whether it qualifies users to analyse and apply the AI systems results appropriately.[69]
As a result of the self-regulation practices which are passed with obligation onto the users in form of an 'instructions for use'. the proposal factually binds and lays the 'reliance of the public sector on the private sector which holds the power over public authorities by setting rules for how to use AI systems'.[70] This is of great concern due to the fact that the vast majority of AI developments and all associated elements, like development platforms, data, knowledge and expertise is dominated by the 'big 5' technology companies in the field.[71]

There was hope the AI Act would contain clear regulations regarding transparency obligations making AI systems more trustworthy for impacted individuals and innovators but this was not the case and transparency issues have been left to self-regulation.[72] Thereby, the AI Act fails to address the fundamental power imbalance between developer and deployer of AI systems and those affected by it.[73]

4. The use of AI and the Consequences for Governmental Institutions, Government Actions and Public Services

With all the proposed changes regarding transparency and classification of AI technology into risk-groups it remains to be seen what effects it will have on certain areas of government

[66] Varošanec, at 9.
[67] Art. 13 (1) AI Act.
[68] Art. 13 (2) AI Act.
[69] Varošanec, at 12.
[70] Varošanec, at 12.
[71] Ulnicane, at 83.
[72] Varošanec, at 17.
[73] Varošanec, at 17.

actions and services. Therefore, examining the effects on Administration Law and certain areas of Criminal Law will be discussed in the following paragraph.

A. AI and day-to-day administrative law processes

Administrative law, as part of public law, is particularly characterized by decisions that are made toward citizens based on state authority. Such decisions may significantly impact people's private and economic life. When it comes to automizing such decisions e.g., by using artificial intelligence in the form of Automatic-Decision-Making, one has to ponder diverse interests. One the one hand, the use of artificial intelligence may render decision-making much more efficient. In this regard, many decisions in the context of administrative law are already standardised to a great extent, e.g., in the case of a municipal authority granting a building permit. Using AI could further increase procedural efficiency and thereby unburden the administration particularly regarding 'standard' decisions. On the other hand, one has to carefully consider the impact of such decisions on the fundamental rights of citizens. Against this background, in many cases, it can be important to assess the decision on a proper case-by-case basis.

In order to safeguard the fundamental rights of citizens in administrative processes, particularly with respect to their procedural judicial rights, a number of procedural principles have evolved over time. One of the core principles is the principle of transparency, meaning an insight into legislative, administrative or judicial proceedings. It is also part of the freedom to expression and the right to information including a right to access official documents. Regarding the use of ADM systems or AI in general administrative procedure, there are certain doubts as to whether or not the procedural rights, particularly the principle of transparency, can be observed. The already mentioned 'black-box' of AI creates in the daily administrative practice a certain asymmetry of information by its nature, due to a lack of transparency towards third parties that in particular do not have access to the AI's algorithm.[74]

B. AI and Criminal Law practice

[74] See also Finck, Michéle 'Automated Decision-Making and Administrative Law' in Cane, P. (ed) *et al., The Oxford Handbook of Comparative Administrative Law* (2020) at 8; Varošanec, at 2.

A world where the crime rates are dropping rapidly and the few crimes that are left are solved with 100 % certainty and in a fair and fast trial – could this be the future? With AI technology, crimes could be predicted, prevented and better and faster investigated, even cancelling out human error during investigations and reducing time-consuming tasks. AI based tools could, for example, detect terrorist propaganda on the Internet, suspicious transactions when selling dangerous or stolen goods or even identify dangerous objects or illegal substances and products.

AI's rapid progress raises challenges regarding benefits and risks for the criminal justice system. The use of AI can cause serious harm if malfunctioning, not only for law enforcement, but also for the courts of law. Fundamental rights must not be undermined and compliance with them must be ensured. Furthermore, discrimination and bias in both the development and use of AI must be prevented. AI should always be monitored by humans.
While criminal law is generally based on the principle that authorities respond to a crime after it has been committed and does not assume that all people are dangerous and must be permanently monitored to prevent potential misconduct, AI carries the risk of being used preemptively. Preventive measures increase the risk of, for example, a surveillance state and thus the violation of fundamental rights. It is necessary for legislators to take action when it comes to how and where AI technology is used in order to ensure legal certainty and respect fundamental rights.

1. The tension between fundamental rights and surveillance

Fundamental rights may be violated when AI is involved in decisions which affect individuals. Related risks not only derive from and regarding investigative proceedings but also with regard to judicial processes. Currently, AI for prevention of crimes is more common in Europe than tools that assist judges.[75]

During investigations concrete powers of intervention must explicitly lay down the basis and limits for the use of AI. This would also make the use of AI comprehensible for the affected parties and promote acceptance among the general public.

[75] Završnik, A. (ed.), *Big Data, crime and social control* (2019), at 194 et seq., lists in detail a series of instruments used by police services in Europe.

Measures could be, for example, surveillance instruments, video and image analytics, facial recognition, mass profiling or the detection of suspicious activities. These could substantially improve crowd surveillance results, as it could help detect patterns and anomalous behavior, even predicting solitary or crowd behavior. AI facial recognition abilities can establish the identity and whereabouts of an individual. Human trafficking, money laundering, fraud and sexual abuse could be detected earlier and even potentially even be prevented, as could criminal networks be uncovered. Though this poses security risks, as criminal law AI can be the target of cyberattacks and can be abused by criminals for malicious purposes. Criminals themselves also take advantage of the progress of AI, as they use AI for their own purposes, for example AI-supported password guessing, hacking and ransomware and cybercrimes in general.

The measures above are always accompanied by interference with fundamental rights as guaranteed in the European Union by the *Charter of Fundamental Rights of the European Union*, the *European Convention on Human Rights*[76] *and the Fundamental Freedoms of the European Union*[77], as well as in the constitutions of the various EU member states.
The use of AI potentially leads to a broader mass of people being affected by the invasive measures, which could result in discrimination based on religious beliefs or political opinions. Especially human dignity, namely the right to the free development of the personality, the right to privacy or protection of personal data are at stake. Surveillance measures always run the risk of restricting other fundamental rights. Individuals could feel restricted in their freedom of expression and/or assembly by foregoing the exercise of their rights out of fear of surveillance. Facial recognition measures can be prone to error, depending on the data and algorithms used to develop the respective AI tool. The impact of such measures on affected parties and the consequences of erroneous evaluations and decisions are incalculable.

2. The AI proposal: The treatment of biometric identification systems and criminal law as high risk

[76] Council of Europe, *European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14*, 4 November 1950, ETS 5, available at: https://www.refworld.org/docid/3ae6b3b04.html.
[77] European Union, *Consolidated version of the Treaty on the Functioning of the European Union*, 26 October 2012, OJ L. 326/47-326/390; 26.10.2012, available at: https://www.refworld.org/docid/52303e8d4.html.

In terms of criminal law, the high-risk systems and the prohibited AI practices should be highlighted. As already mentioned, real-time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement are prohibited by the AI Act. But first, there needs to be a distinction made between 'real time' biometric systems, in which the system runs the process (nearly) live and instantaneously[78] and 'post' biometric systems for law enforcement purposes in public spaces[79]. Prohibited as a matter of principle is only the use of 'real time' biometric systems[80], except for three situations where its use is deemed justified for reasons of substantial public interest:

1. Search for victims of crime, including missing children
2. Threat to life or physical integrity or of terrorism
3. Serious crime (EU arrest warrant)[81]

The reason for the distinction between the two types of biometric systems, is that, 'real time' biometric systems are 'considered particularly intrusive in the rights and freedoms of the concerned persons, to the extent that it may affect the private life of a large part of the population, evoke a feeling of constant surveillance and indirectly dissuade the exercise of the freedom of assembly and other fundamental rights'[82] In a first reaction, the European civil rights organization (EDRi) called on the Parliament to improve the Commission's proposals. EDRi criticized, among other things, that they did not go far enough in excluding 'biometric mass surveillance', while the exceptions were too far-reaching.[83] The demand was also supported by 116 members of the European Parliament in an open letter to President von der Leyen.[84] Meanwhile, more than 71,000 European citizens have signed a petition for a ban on

[78] Art. 3 (37) AI Act.
[79] Art. 3 (38) AI Act.
[80] Art. 5 (d) AI Act.
[81] Art. 5 (d) i-iii; Recital 19 AI Act.
[82] Recital 18 AI Act.
[83] EdRi, *Open letter: Civil society call for the introduction of red lines in the upcoming European Commission proposal on Artificial Intelligence* (2021), available at https://edri.org/wp-content/uploads/2021/01/EDRi-open-letter-AI-red-lines.pdf; see also EdRi, *Seeking your support for a specific ban on biometric mass surveillance practices on fundamental rights grounds* (2021), available at https://edri.org/wp-content/uploads/2021/04/Letter-from-51-civil-society-organisations-seeking-your-support-for-a-ban-on-biometric-mass- surveillance-practices.pdf
[84] EdRi, *Open Letter* (2020), available at https://edri.org/wp-content/uploads/2021/03/MEP-Letter-on-AI-and-fundamental-rights-1.pdf.

biometric mass surveillance practices as part of the 'Reclaim Your Face' campaign, and the number continues to grow.[85]

Accordingly, 'post' and also 'real time' remote biometric identification systems are classified as high-risk.[86] Furthermore law enforcement falls under high risk.[87] The European Parliament acknowledges that 'given the role and responsibility of police and judicial authorities and the impact of decisions they take for the purpose of preventing, investigating, detecting or prosecuting crime or enforcing criminal sanctions, the use of AI applications must be considered high-risk when there is a possibility that it will have a significant impact on the lives of individuals'[88]

5. Conclusion

Transparency is a central and significant concept in law and important when creating trust, making it challenging to find the right balance between established principles and new technology that undergoes complicated development processes with numerous people involved.

The main concern is the difficulty for individuals to challenge automated decisions, especially when the decision-making process and the data used as a basis for the algorithms involved are not clearly stated and publicly available. Having uncertain regulations regarding the risk-assessment could have major consequences especially when AI technology is used in judiciary processes, for example assisting judges in their decision-making process and improving the foreseeability of the use of law and the coherence of judicial decisions. Therefore it is crucial to safeguard the right to the lawful judge, as well as to respect the fair trial principle. The use of sensitive data[89] in judgments, carries the risk of discrimination and

[85] European Citizens' Initiative (ECI) 'Reclaim Your Face' campaign, as of June 14th, https://reclaimyourface.eu/ (last accessed 14.06.2022).
[86] Recital 33 AI Act.
[87] Art. 6 (2) in conjunction with Annex III 6 AI Act.
[88] EP Report *On artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters* (2020/2016(INI)) (2020), available at www.europarl.europa.eu/doceo/document/A-9-2021-0232_DE.html.
[89] Such as racial or ethnic origin, political opinions, religious or political beliefs, socioeconomic conditions, or data on health or sexual orientation.

bias which needs to be avoided. Therefore, government measures using AI based technology should be just as transparent as a human decision.[90]

Due to the socio-technical complexity of AI technology and ADM systems and the multidisciplinary character linking data acquisition, hardware and software architectures like algorithms, and the skills, mindset and judgments of designers, coders, commercial providers and civil servants [91] accomplishing adequate transparency and satisfying the needs and expectations of all stakeholders is very challenging.[92] In order to create trustworthy automated AI decision-making systems for legal backgrounds, interdisciplinary collaboration between software programmers/engineers and legal experts is essential to reduce legal errors and ensure sufficient explanation of how the decision was reached.[93] The complex design decisions regarding legal issues would need to be made together with experts and developers in advance. Creating a legal framework that finds the balance between established legal concepts, like transparency and new technology is never easy and usually leads to very complex regulations, not always satisfying all stakeholders and sometimes even creating new problems without solving existing ones.

Overall, there is the technical possibility to create sufficient transparency. Therefore, it should not be accepted that AI technology is guided only by efficiency and thus commercial motives.

There are many options of regulations for AI based technology and the EU´s risk-based approach is one step in the right direction for individual solutions but it does not solve all the problems just yet.
One of the main concerns is that it leaves transparency issues to self-regulation and does not provide safeguards concerning the review of new AI technology, leaving it to developers to choose, acquire and use data, algorithms are based upon, which leaves a risk of abuse in a competitive field of the global market.[94] This creates not only dependency but leads to a power

[90] Timmers, *supra* note 88, at 51.
[91] Timmers, Paul 'AI Challenging Sovereignty and Democracy' 20 *4* TPQ (2021/22) 45, at 51.
[92] *Ibid.*
[93] Mökander, J., Axente, M., Casolari, F., Luciano Floridi and J Mökander, *Conformity Assessments and Post-Market Monitoring: A Guide to the Role of Auditing in the Proposed European AI Regulation* (2021), available at https://doi.org/10.1007/s11023-021-09577-4.
[94] Varošanec, at 17.

imbalance between AI developers and public authorities and other bodies using such technology.[95]

As pointed out by some critics, the AI Act does not provide for a general fundamental rights risk assessment for all AI technology but only for those now classified as 'high risk'.[96] The criteria for why certain AI technology are classified in certain risk-groups is not justified by externally reviewable criteria making it difficult to add new future systems according to the criteria in Art. 7 AI Act.[97]

The aim and goal of the European Commission is to make AI secure and trustworthy. The proposed rules will have to be implemented and enforced by Member States. In addition, a European Artificial Intelligence Board (EAIB), consisting of representatives from every Member State, to help the Commission decide which AI systems count as "high-risk" and to recommend changes to prohibitions, will be established. [98]

The law must and should move ahead with times preparing for new challenges created by advancing technologies. The AI Act is a first step in the right direction and could set a global standard for regulating AI technology internationally.

[95] *Ibid.*
[96] Edwards, at 5.
[97] *Ibid.*
[98] Art. 46 AI Act.