

The challenge of defining artificial intelligence in the EU AI Act

Received (in revised form): 13th November, 2023



Theodore S. Boone

Member of the Faculty, Corvinus University Budapest and Of Counsel, Dentons, Hungary

Theodore Boone is a member of the faculty at Corvinus University in Budapest and Of Counsel in the Budapest office of Dentons. His work focuses on artificial intelligence and other information technology matters. At Corvinus University, Mr Boone teaches a course titled 'Business Leadership in the Age of Artificial Intelligence', and courses on international business law and co-teaches a course on EU economic policy and institutions. In his work, Mr Boone draws on his experience arising from previous leadership positions in the US and Europe at international law firms and one of the Big 4. He is a former president of the American Chamber of Commerce in Hungary and a former adjunct professor at Georgetown University's School of Law in Washington, DC. Mr Boone is a member of the Washington, DC, New York and Supreme Court of the United States bars. He received his JD from Columbia University School of Law and also holds an LLM from Hungary's Eötvös Loránd University School of Law.

Department of Business Law, Corvinus University of Budapest, Fővám tér 8, 1093 Budapest, Hungary
Tel: +36-70-458-5203, E-mail: theodore.boone@uni-corvinus.hu

Abstract The EU Commission, the EU Council and the EU Parliament have each issued their own versions of the text of a new EU AI Act. Throughout the gestation of the EU AI Act a core and complex question has arisen: how should 'AI system' be defined in the EU AI Act? This paper examines the evolution of the definition of 'AI system' in the draft EU AI Act. This paper suggests that, in order to achieve the EU's goals, a definition of 'AI system' which is clear and cannot be modified outside the EU legislative process, which is sufficiently broad to accommodate future technological developments and which focuses on systems which make predictions, recommendations and decisions would appear to be both the most practical and the most appropriate.

KEYWORDS: AI, artificial intelligence, AI system, EU, EU AI Act, regulation

WHAT IS ARTIFICIAL INTELLIGENCE? Origins of artificial intelligence

The concept of artificial intelligence (AI) is by no means new. Alan Turing examined the question 'can machines think?' in his seminal 1950 article 'Computing Machinery and Intelligence'.¹ The proposal for the 1956 'Dartmouth Summer Research Project on Artificial Intelligence' is widely credited with first expressly using the term 'artificial intelligence'. The 31st August, 1955 proposal for the conference stated:

We propose that a 2-month, 10-man study of artificial intelligence be carried out

during the summer of 1956 at Dartmouth College in Hanover, New Hampshire.

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.²

Between the time of the 1956 'Dartmouth Summer Research Project on Artificial Intelligence' and the mid-1970s focus on and funding for AI research was substantial. However, lack of sufficient computing power inhibited continued advancement. In recent decades computing

power has increased exponentially. This fact, combined with the relatively recent vast increase in the availability of data, which can be used to train AI systems, and the expanded abilities and use of cloud computing, has once again brought AI's development and abilities to the forefront.³

As an example of these developments one can consider AI advancements related to the Chinese board game Go. This game has more potential moves than known atoms in the universe. In 2017 Google's DeepMind introduced AlphaGo Zero. This system learned how to play Go by being provided with nothing other than the rules of Go and then playing against itself, starting from completely random play. By teaching itself in this manner, AlphaGo Zero developed its own style of creative play and was soon able to play at a higher level than a Go world champion.⁴ As stated by DeepMind:

AlphaGo Zero [. . .] learns to play simply by playing games against itself, starting from completely random play. In doing so, it quickly surpassed human level of play [. . .]. It is able to do this by using a novel form of reinforcement learning, in which AlphaGo Zero becomes its own teacher. The system starts off with a neural network that knows nothing about the game of Go. It then plays games against itself, by combining this neural network with a powerful search algorithm. As it plays, the neural network is tuned and updated to predict moves, as well as the eventual winner of the games. [. . .] After just three days of self-play training, AlphaGo Zero emphatically defeated the previously published version of AlphaGo – which had itself defeated 18-time world champion Lee Sedol – by 100 games to 0.⁵

Since 2017's AlphaGo Zero the power and abilities of such AI systems have only increased. This power and these abilities pierced the broader public consciousness in late 2022 when Open AI released ChatGPT, a so-called large language model tool that can generate

sophisticated human-like text responses to queries and requests, to the public.⁶

As a result of these recent developments the terms 'artificial intelligence' and 'AI' have now entered the everyday lexicon, but generally without much in the way of specificity. Does the definition of AI include a software system that is making a medical diagnosis and predicting the outcome of a course of treatment? Should a definition of AI include a software system that is making recommendations as to which applicants to admit to a university and predicting how well they will do if admitted? Does a software system that is making sentencing recommendations to a judge in a judicial procedure fall within its definition? Should everyday applications such as a music app that makes recommendations as to songs you might like or a shopping app that makes recommendations as to items you may wish to purchase be considered AI? In the last few years, the EU, in an effort to regulate AI, has been confronting these very questions.

The European Commission proposes an EU AI Act

On 21st April, 2021, the European Commission issued a proposal for the regulation of AI by the EU (the 'European Commission Proposal').⁷ This proposal was a significant step in an ongoing focus by the EU over several years on AI related matters.⁸ The proposal included a draft text of an EU Act governing AI. This Act as then proposed would govern the use of AI not only by entities based in the EU but also the use of AI by entities regardless of their geographic base around the globe doing business in the EU.

Given that the proposed EU AI Act seeks to regulate the use of AI, the setting forth in the proposed EU AI Act of a definition of what actually constitutes AI is of fundamental importance. However, as the proposed EU AI Act has moved forward in the EU legislative process it has proven to be no easy task for the relevant parties within the EU to agree on a definition of AI. Indeed,

disagreement on the definition of what should constitute AI is one of the core factors inhibiting the finalisation of an EU AI Act.

When examining the question of how AI should be defined by the EU it is important to keep in mind that what is generally referred to as AI contains numerous subfields. Perhaps the most well-known subfield is referred to as machine learning, that is a system which can extract patterns and rules from data. However, there are other subfields or overlapping fields as well, including natural language processing (systems having to communicate in the world’s natural languages) and machine perception (systems that analyse and draw deductions from visual, auditory and other sensors).

THE CORE OBJECTIVES OF THE EU AI ACT

When issuing the European Commission Proposal and the draft EU AI Act text included therein, the European Commission set forth the following as its core objectives:

- a) ensuring that AI systems placed on the EU market and used are safe and respect existing law on fundamental rights and EU values;
- b) ensuring legal certainty to facilitate investment and innovation in AI;
- c) enhancing governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems;

- d) facilitating the development of a single market for lawful, safe and trustworthy AI applications and preventing market fragmentation;⁹ and
- e) putting in place a regulation that is flexible and as ‘future proof’ as possible given the fast pace of technological developments.¹⁰

One of the challenges faced by EU legislators in the context of seeking to come to an agreement on a definition of AI is that these five objectives may conflict with one another in certain circumstances.

Of these objectives one that is perhaps particularly important relates to the issue of developing trustworthy AI. If the EU AI Act does not foster an environment where individuals trust an AI system the EU AI Act will have failed in developing a legal framework for AI where EU fundamental rights and values are respected. Closely related to this issue of trust is the question of transparency. In this context transparency is, in its essence, the ability of an AI user to understand the grounds and reasoning behind an AI generated prediction or recommendation. That is, for transparency to exist a user should be able to connect the dots back from D (the prediction or recommendation) to C to B to A (the query). In cases of an AI generated medical diagnosis, for example, a user would need to be able to understand the basis of, and reasoning behind, such diagnosis. Fostering such transparency in the cases of AI systems which have in essence

Objectives of EU AI Act:

- Promoting safety and respect for existing law and values
- Ensuring legal certainty to facilitate investment and innovation
- Enhancing governance and effective enforcement of existing law
- Developing of a single market for lawful, safe and trustworthy AI
- Creating a regulation that is flexible and “future proof”

Figure 1: Objectives of the EU AI Act

taught themselves, such as the AlphaGo Zero system mentioned above, is particularly difficult since the basis for any individual AI system determination may not be readily apparent and therefore not readily transparent (Figure 1).

THE EUROPEAN COMMISSION APPROACH

Risk based

The draft EU AI Act text in the European Commission Proposal took a risk-based approach to the regulation of AI. In essence, the level of risk posed by an AI system was established as the central factor in determining whether an AI system would be entirely prohibited, subject to substantial regulations and regulatory oversight or subject to minimal rules and requirements.

In the draft EU AI Act text in the European Commission Proposal, AI systems were placed into three categories: (i) those that present ‘unacceptable risk’; (ii) those that present ‘high risk’; and (iii) those that present ‘low risk’.

AI systems that present an unacceptable risk were described as those that contravene fundamental EU values.¹¹ Examples of AI systems presenting an unacceptable risk included certain systems which deploy subliminal techniques to distort a person’s consciousness, certain social scoring systems and, except for limited circumstances, real-time biometric systems deployed in publicly

accessible spaces for the purpose of law enforcement.¹²

AI systems that were not entirely prohibited by virtue of presenting an unacceptable risk but designated in the draft EU AI Act text included in the European Commission Proposal as ‘high risk’ were those that pose a significant risk to the health, safety or fundamental rights of persons.¹³

The draft EU AI Act text included in the European Commission Proposal also contained an enforcement mechanism which included the potential for fines of up to €30 million or 6 per cent of the total worldwide annual turnover of the violating entity for the year preceding the violation, whichever is higher (Figure 2).¹⁴

Scope of application

The draft EU AI Act text set forth in the European Commission Proposal provided that the EU AI Act would apply to:

- (i) providers placing on the market or putting into service AI systems in the EU, irrespective of whether those providers are established within the EU or in a third country;
- (ii) users of AI systems located within the EU; and
- (iii) providers and users of AI systems that are located in a third country, where the output produced by the system is used in the EU.¹⁵

EU’s risk
-based
approach:

Prohibited AI systems

High-risk AI systems

Low-risk AI systems

Figure 2: The EU’s risk-based approach

EU AI Act
would
apply to:

Providers placing on the market or putting into service AI systems in the EU

Users of AI systems located within the EU

Providers and users of AI systems that are located in a third country, where the output produced by the system is used in the EU

Figure 3: Who the EU AI Act would apply to

Hence, the regulations as set forth in the draft of the EU AI Act contained in the European Commission Proposal were considered to have a global reach, as any entity, regardless of geographical seat, seeking to place an AI system on the market within the EU or using in the EU output produced by an AI system would be required to comply with its terms (Figure 3).

THE DEFINITION OF AI SYSTEM AS SET FORTH IN THE EU COMMISSION PROPOSAL

Wording

The draft EU AI Act contained in the European Commission Proposal defined AI in the context of defining an ‘AI system’. There is no definition of AI in the draft separate from the definition of AI system. AI system is defined in that draft as follows:

‘artificial intelligence system’ (AI system) means 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.¹⁶

Annex I

This definition focused on the functionality of the software and referred to software

developed by techniques and approaches listed in an annex, Annex I, to the EU AI Act. This Annex I itself set forth such techniques and approaches as follows:

- (a) machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;
- (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 optimisation methods.¹⁷

Amendments to Annex I

The draft EU AI Act set forth in the European Commission Proposal stated that the European Commission could amend the list of techniques and approaches listed in the referenced Annex I in order to update the list to address market and technological developments on the basis of characteristics that are similar to the techniques and approaches listed in Annex I at the time of the enactment of the EU AI Act.¹⁸

The draft EU AI Act set forth in the European Commission Proposal therefore delegated substantial authority to the European Commission to adjust the type of software that is subject to the EU AI Act

after the EU AI Act's enactment and entry into force. Such a delegation enhanced the ability of the EU AI Act to adapt to evolving technologies without the need for an amendment to the EU AI Act itself. Such a delegation therefore served to promote the EU AI Act's goal of a flexible and 'future proof' framework. However, such a delegation also granted a significant power to the European Commission, as the European Commission would then have been entitled to adjust the software types which are subject to the EU AI Act without the need for the more expansive process that would have been required to amend, through the legislative process, the EU AI Act itself.

Human defined objectives

The definition of AI system set forth in the European Commission Proposal's draft EU AI Act also specified that the software at issue must be able to generate outputs 'for a given set of human-defined objectives'.¹⁹ The phrase 'for a given set of human-defined objectives' was not itself defined in the draft of the EU AI Act contained in the European Commission Proposal. A reasonable interpretation of the phrase could be that the software must be generating outputs based on targets established and specified by a natural person in order to achieve a specified outcome. Such targets could be, by way of only a few examples, establishing a credit score, identifying indications of cancer on X-Rays or weather forecasting.

Outputs

The definition of AI system set forth in the draft EU AI Act contained in the European Commission Proposal then referred to the software generating outputs 'such as content, predictions, recommendations, or decisions'.²⁰ As a result of the use of the words 'such as' in this phrase, the outputs listed, that is content, predictions, recommendations and decisions, may be read

only as examples. Hence, software outputs of any form could have been potentially subject to the definition.

Influencing environments

The final portion of the definition of AI system set forth in the draft EU AI Act contained in the European Commission Proposal required that the outputs are ones that were 'influencing the environments they interact with'.²¹ The Preamble to the draft EU AI Act contained in the European Commission Proposal specified that such environments could be either physical or digital.²² The Preamble to the draft also clarified that the AI systems subject to regulation could be ones that operate with varying levels of autonomy and could be used on a stand-alone basis or as a component of a product, irrespective of whether the system is physically integrated into a product or served the functionality of a product without being integrated therein.²³

Scope of definition

The definition of AI system set forth in the draft EU AI Act forming part of the European Commission Proposal was broad. The definition created the potential for the EU AI Act to cover more traditional software in a manner that could inhibit one of the draft's stated goals — economic development and innovation. In addition, by delegating significant authority to the European Commission via its enablement of the European Commission to amend the techniques and approaches listed in the Annex I referred to in the definition and used to develop the software at issue the draft, arguably at least, promoted its stated goal of being 'future proof'. However, such a delegation of authority to the European Commission accomplished that promotion in a manner that created less legal certainty and in a manner that

European
Commission
proposed
definition of
AI System:

Software that is developed with one or more of the
techniques and approaches listed in an annex

That can, for a given set of human-defined objectives,
generate outputs such as content, predictions,
recommendations or decisions influencing the
environments they interact with

Figure 4: The European Commission's proposed definition of an AI system

allowed the EU AI Act as presented in the European Commission Proposal to be amended outside the rigours and debate of an actual EU legislative amendment to the Act (Figure 4).

THE EU COUNCIL'S 25TH NOVEMBER, 2022 COMPROMISE PROPOSAL

New definition of AI system

The EU Council began examination of the European Commission Proposal soon after the proposal's release. Following that examination a compromise text of the EU AI Act was released by the EU Council on 25th November, 2022 (the 'EU Council Proposal').²⁴

The EU Council Proposal substantially narrowed the definition of 'AI system' from that which was contained in the European Commission Proposal.

The definition of AI system as set forth in the EU Council Proposal was as follows:

(1) 'artificial intelligence system' (AI system) means a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge-based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts.²⁵

Goal of distinguishing from simpler systems

In the EU Council's commentary on the EU Council Proposal a stated goal of the modification of the definition of AI system from that which was contained in the European Commission Proposal was to distinguish AI from 'more classical software systems' and to narrow down the definition to systems 'developed through machine learning approaches and logic- and knowledge-based approaches'.²⁶

Similarly, in the Recitals to the EU AI Act as set forth in the EU Council Proposal it was stated that the definition of AI system should be based on key functional characteristics of artificial intelligence such as its learning, reasoning or modelling capabilities, distinguishing it from simpler software systems and programming approaches'.²⁷ In this regard one of the Recitals to the EU AI Act contained in the EU Council Proposal suggested that that software which undertakes a mechanical execution of operations should be excluded from the definition of AI system and therefore the coverage of the EU AI Act when the Recital stated 'A system that uses rules defined solely by natural persons to automatically execute operations should not be considered an AI system'.²⁸

New concepts

In its effort to distinguish an AI system from other software systems and to narrow its scope, the EU Council Proposal also introduced several new concepts.

Elements of autonomy

First, the definition in the EU Council Proposal referred to AI systems ‘designed to operate with elements of autonomy’. What constituted ‘elements’ and ‘elements of autonomy’ were themselves however not defined in the EU Council Proposal. In this context, Recital 6 of the EU Council Proposal stated, in a not entirely illuminating fashion, that ‘The concept of the autonomy of an AI system relates to the degree to which such a system functions without human involvement’.²⁹ The degree to which such a system would need to function without human involvement that would be sufficient to fall within the concept of ‘elements of autonomy’ was however not described.

Human involvement in objectives/ data and inputs

Secondly, the concept of ‘for a given set of human defined objectives’ which was set forth in the European Commission Proposal was removed from the definition of AI system in the EU Council Proposal. In its place was set forth the concept of ‘based on machine and/or human-provided data and inputs’.

Hence, in the definition of AI system itself set forth in the EU Council Proposal it was no longer a requirement that the content generated by the AI system be associated with goals established by a human. Rather, under the phraseology of the definition of AI system contained in the EU Council Proposal, the content generated must be based merely on data or inputs provided by either a human or a machine.

Therefore, under this definition, the AI system would be one that relies on data and inputs that are either generated by machines, such as sensors, cameras or algorithms, or provided by humans, such as text, images or voice. Presumably, the data and inputs may be structured or unstructured and static or dynamic and may come from various sources. The definition does not discuss issues related to the quality, quantity or origin of the data and

inputs. Failure to address issues related to the quality of data and inputs is potentially perilous as such failure could result in bias predictions and recommendations. For example, an AI system could be asked to recommend the qualities that make an excellent engineer and in such a context be provided with data and inputs related to the qualifications, experience, publications and awards of engineers. If such data as a whole show that the majority of engineers are male the AI system could determine that being male is a preferred quality in an engineer, a determination which is obviously false and biased.

The role of the involvement of human-defined objectives was not completely abandoned in the EU Council Proposal, however. Although the definition of AI system itself in such proposal no longer referred to human-defined objectives, the Recitals to the EU AI Act contained in the EU Council Proposal did state that the AI system is one that infers a way to achieve a set of ‘final objectives given to them by humans’.³⁰

Inference using machine learning and/or logic and knowledge-based approaches

As a third new concept, the definition of AI system incorporated into the EU Council Proposal stated that the system must be one that ‘infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge-based approaches’. This phrase in effect replaced that portion of the definition of AI system contained in the draft EU AI Act included in the European Commission Proposal, which referred to software developed ‘with one or more of the techniques and approaches listed in Annex I’.³¹

Machine learning

The Recitals to the EU Council Proposal provide a useful description of machine learning:

Machine learning approaches focus on the development of systems capable

of learning and inferring from data to solve an application problem without being explicitly programmed with a set of step-by-step instructions from input to output. Learning refers to the computational process of optimising from data the parameters of the model, which is a mathematical construct generating an output based on input data.³²

Hence, the concept of machine learning in the definition is tied to the notion of the ability to solve a problem without step-by-step instructions as to how to do so.

The EU Council Proposal also stated:

The range of problems addressed by machine learning typically involves tasks for which other approaches fail, either because there is no suitable formalisation of the problem, or because the resolution of the problem is intractable with non-learning approaches.³³

This linkage of the problems often addressed by machine learning to those ‘for which other approaches fail’ due to there being no suitable formalisation of the problem or due to intractability of the problem was perhaps a questionable way to describe machine learning. It appears that there are certainly problems where machine learning can be used even though there are other methods to address the problem. As one example, the use of machine learning for diagnostic purposes in the health care sector might be considered. For example, in Hungary, machine learning has been used to assist medical professionals in reading an X-Ray and in making a diagnosis related thereto.³⁴ In such a case the use of machine learning could be viewed not so much as to address a problem because there is no other method to do so but rather to augment efficiencies by acting as a second or third reader of an X-Ray, a task that might have previously been performed by a human.

The role of inference

The definition of AI system contained in the EU Council Proposal introduced an entirely new concept absent from the definition of AI system contained in the European Commission Proposal. This new concept was the concept of the AI system ‘inferring’ how to achieve a given objective. The verb ‘to infer’ is itself not defined, however. A reasonable interpretation might be that the AI system uses the referenced approaches to learn from data and inputs, to reason, predict, synthesise, recommend or solve problems or to achieve a predefined or self-defined goal or outcome. The AI system may use different types of AI approaches, such as supervised, unsupervised or reinforcement learning, or logic and knowledge-based systems, such as expert systems, ontologies or rules. There are various definitions of the verb ‘to infer’. For example, ‘infer’ can mean (i) to derive as a conclusion from facts or premises; (ii) to guess or surmise; (iii) to involve as a normal outcome of thought; (iv) to point out or indicate; and (v) to suggest or hint.³⁵ Given the range of meaning available to the concept of ‘infer’, it was unfortunate that a definition of ‘to infer’ was not included in the draft of the EU AI Act contained in the EU Council Proposal. However, of the options of the definition of ‘infer’, it would appear that that which would be most appropriate in the context of the definition of AI system appearing in the EU Council Proposal would be the concept of ‘to derive as a conclusion from facts or premises’, as such a definition would appear to be most compatible with the concept of an inference being drawn through the use of machine learning and/or logic and knowledge-based approaches.

Role of the European Commission post enactment

In the EU Council Proposal the EU Council dispensed with the concept of

EU Council
proposed
definition of
AI system:

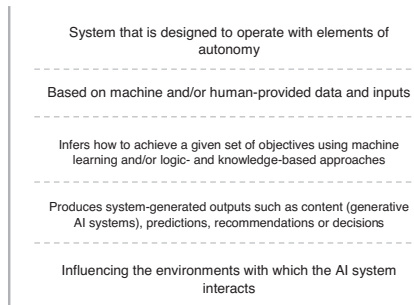


Figure 5: The EU Council's proposed definition of an AI system

enabling the European Commission to in effect modify the definition of AI system by modifying the Annex I list of techniques and approaches attached to the European Commission Proposal's draft of the EU AI Act by delegated acts. Such an approach eliminated the possibility that a core definition in the EU AI Act, that of what constitutes an AI system, could be altered by the European Commission acting outside the full EU legislative process required for an amendment to an Act. In this sense the elimination of the European Commission's powers to amend the definition of AI system provided more certainty and predictability. However, the elimination of the European Commission's powers to amend the definition of AI system also potentially reduced the ability of the EU AI Act to fulfil its goal of being as future proof as possible in the fast-changing environment of information technology (Figure 5).

THE EUROPEAN PARLIAMENT'S 14TH JUNE, 2023 PROPOSED TEXT

Since the release of the EU Council Proposal committees of the European Parliament have focused on matters related to the EU AI Act, including the appropriate definition of AI systems. On 9th May, 2023 the European Parliament's Committee on the Internal Market and Consumer Protection and the European Parliament's Committee on Civil Liberties,

Justice and Home Affairs jointly issued a draft text of the EU AI Act.³⁶ Then, on 14th June, 2023, the European Parliament adopted a proposed text of the EU AI Act (the 'European Parliament Proposal').³⁷ An entirely new definition of AI system appears in the European Parliament Proposal. This definition is as follows:

(1) 'artificial intelligence system' (AI system) means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments.³⁸

In establishing this definition, the European Parliament agreed on a definition of AI systems that essentially tracks the definition of AI system set forth in the 'Recommendation of the Council on Artificial Intelligence' of the Organisation for Economic Cooperation and Development (OECD) and adopted by the OECD Council at ministerial level on 22nd May, 2019.³⁹ In the OECD 2019 AI recommendation, the OECD defined an AI system as:

a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.⁴⁰

Here, of note is the fact that the US National Artificial Intelligence Initiative Act of 2020 defines ‘Artificial Intelligence’ in a manner which also tracks the definition of AI system set forth in the OECD 2019 AI Recommendation. The definition of ‘Artificial Intelligence’ in the US National Artificial Intelligence Initiative Act of 2020 is, ‘a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments’.⁴¹

The US National Institute of Standards and Technology, in its January 2023 publication setting forth a risk management framework related to AI, specifically states that it adapted its definition of AI system from the OECD 2019 definition of AI system. The US National Institute of Standards and Technology definition of AI system and the OECD 2019 definition of AI system are nearly the same, word for word. The US National Institute of Standards and Technology definition refers to an AI system as follows:

as an engineered or machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.⁴²

In addition, the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence issued by

The White House on 30th October, 2023, which, among other matters, instructs US Federal government agencies to take a broad set of actions relating to AI, uses the definition of AI set forth in the US National Artificial Intelligence Initiative Act of 2020, a definition which, as noted above, tracks the definition of AI system set forth in the OECD 2019 AI Recommendation.⁴³

Therefore, the definition of AI system used by the EU Parliament essentially tracks the definition of AI system set forth in the OECD 2019 AI Recommendation, the US National Artificial Intelligence Initiative Act of 2020, the US National Institute of Standards and Technology’s publication setting forth a risk management framework related to AI, and the White House AI Executive Order’s definition of AI or AI system (Figure 6).

This commonality of definition as between the EU Parliament, US regulatory initiatives and the OECD’s 2019 definition of AI system appeared to indicate that EU and US regulatory initiatives were moving in the direction of regulations that would largely be in harmony as concerned the definition of AI system. However, on 8th November, 2023, the OECD issued an updated definition of AI system. This updated OECD definition of AI system is as follows:

An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions,

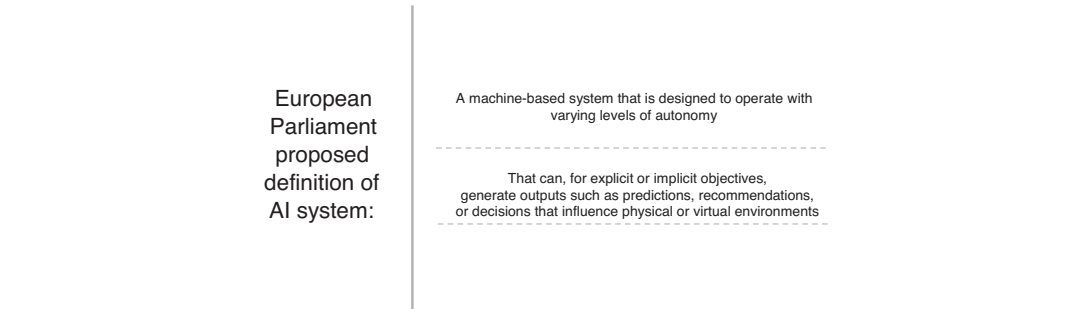


Figure 6: The European Parliament’s proposed definition of an AI system

content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.⁴⁴

The OECD 2023 definition of AI system made several material adjustments to the OECD 2019 definition of AI system. These adjustments to the OECD 2019 definition include: (a) expanding the description of AI system outputs to include, not only predictions, recommendations and decisions, but also the broad term of ‘content’; (b) removing the requirement that objectives must be ‘human defined’; (c) adding the concept that an AI system be one which ‘infers’; (e) adding the words ‘such as’ before the list of outputs generated by AI systems (making the list one of examples rather than an exclusive list); and (f) inserting the concept that objectives may be either ‘explicit or implicit’. Interestingly, each of these recent adjustments by the OECD to the OECD’s definition of AI system circle back to components of each of the definitions of AI system issued by the EU Commission, the EU Council and the EU Parliament issued prior to the OECD’s issuance of OECD 2023 definition of AI system. For example, the EU Commission definition of AI system referred to outputs ‘*such as content, predictions, recommendations, or decisions*’.⁴⁵ The EU Council’s definition of AI system removed the EU Commission’s reference to ‘human defined objectives’ and replaced it with ‘*machine and/or human-provided data and inputs*’. The EU Council also introduced the concept of an AI system being one which ‘infers’.⁴⁶ The EU Parliament introduced the concept that objectives may be ‘explicit or implicit’.⁴⁷

Given that, to date, the EU Parliament and US regulatory initiatives have largely followed the lead of the OECD as to the definition of AI system, it is certainly possible that the EU Parliament and US

regulatory initiatives will continue to follow such OECD lead in the future, including in the context of any final version of the EU AI Act and implementation by US Federal government agencies of the directives issued by the White House in the White House AI Executive Order.

CONCLUSIONS

The definition of AI system, that which is being regulated, is clearly of fundamental importance to the process of the finalisation of the EU AI Act. The fact that each of the central bodies of the EU involved in the EU AI Act legislative process, the European Commission, the EU Council and the European Parliament, has proposed a distinct version of this definition demonstrates both the importance placed on establishing a proper definition and the difficulty in doing so.

The European Commission’s definition of AI system focused on the functionality of the software and referred to software developed by techniques and approaches listed in an annex to the EU AI Act. The European Commission’s proposal provided that the European Commission could amend the list of techniques and approaches listed in the annex. Such an approach augmented the ability of the EU AI Act to adapt to evolving technologies without the need for an amendment to the EU AI Act itself and fostered a ‘future proof’ legislative framework. Under such an approach the European Commission would be able to amend the software types that are subject to the EU AI Act without a need for an amendment to the EU AI Act itself.

The EU Council’s proposed draft of the EU AI Act, which followed the European Commission’s proposal, narrowed the European Commission’s definition of AI system. In narrowing the definition of AI system the EU Council took the position that the definition should be based on key functional characteristics of artificial

intelligence such as its learning, reasoning or modelling capabilities in order to separate the definition of AI system from simpler software systems.

The European Parliament's proposed definition of AI system, which followed the proposals issued by the European Commission and the EU Council, essentially tracks the OECD's 2019 definition of AI system as well as the definitions of AI or AI system set forth in the US National Artificial Intelligence Initiative Act of 2020, the US National Institute of Standards and Technology's publication which set forth a risk management framework related to AI and the White House AI Executive Order. A benefit of the EU Parliament's approach to the definition of AI system, as opposed to that of the EU Commission and the EU Council, is that if the EU Parliament's approach is adopted in the final version of the EU AI Act, such an approach would largely be in harmony with US AI regulatory initiatives to date.

The use of the European Parliament's definition of AI system in the EU AI Act would have several other benefits as well compared to the definitions of AI system proposed by the European Commission and the EU Council.

For example, the use of the European Parliament's definition of AI system in the EU AI Act would not, unlike the European Commission's definition of AI system, allow for its modification by the European Commission. It could be argued that an inability under the EU AI Act of the European Commission to amend the definition of AI system is a detriment rather than a benefit because the EU AI Act would be less 'future proof'. However, the European Parliament's definition, due to its broad wording, in and of itself does provide for significant forward-looking flexibility which could be relevant in the context of developing technologies.

Secondly, unlike the European Commission's and the EU Council's

definitions of AI system, the European Parliament's definition is relatively clear. In essence, under such a definition, an AI system would consist of a system with the following three core characteristics: (i) it is machine-based; (ii) it uses human-defined objectives; and (iii) it makes predictions, recommendations or decisions. Clarity of the definition of AI system in any EU AI Act is of critical importance. A lack of clarity will inhibit, rather than work to fulfil, the EU-stated goals of the EU AI Act, such as protecting fundamental rights and facilitating innovation.

Of these three core characteristics, that which is perhaps most crucial, and most useful in drawing a distinction between an AI system and other software systems, is the requirement that the system be one which makes predictions, recommendations or decisions. These three concepts, predictions, recommendations and decisions, are relatively clear and straightforward to interpret and also address the area of the use of such systems which are likely to be of greatest concern in the context of the protection of the fundamental rights of EU citizens. Software systems which do not make predictions, recommendations or decisions could be viewed simply as tools providing a form of mechanical assistance and inappropriate for regulation under the EU AI Act. However, software systems which are making predictions, recommendations or decisions appear to be those which, if left unregulated, would pose the greatest potential threat to the fundamental rights of EU citizens. As examples of systems which could post such threats to fundamental rights if unregulated, one can consider a system which predicts the success at a university of an applicant to such university, a system which recommends the length of a prison sentence for an individual convicted of a crime or a system which decides whether an applicant for a loan should receive such loan. In this context

it is reasonable to question whether, as both the EU Council proposed and the OECD 2023 definition of AI system have done, expanding the definition of AI system to include not only systems which generate predictions, recommendations and decisions but also systems which, in a now non-exclusive ‘such as’ list, generate ‘content’ is prudent as, under such a definition, it would appear that potentially any form of output at all would fall within the definition. That is, one must ask the questions: What meaningful set of systems would not fall within the definition? Would a definition that is so broad stifle innovation and economic development?

Although not as broad as the OECD 2023 definition of AI system, the European Parliament’s definition of AI system is already a broad one. Notwithstanding that this is the case, one must keep in mind that were the EU Parliament’s definition, or even the broader OECD 2023 definition, to be used in an EU AI Act as currently structured, not all systems falling within the definition of AI systems would be subject to the same degree of regulation under the EU AI Act.

Under the current structure of the EU AI Act, software systems falling within the definition of AI systems would be subject to different rules and levels of oversight depending on whether they are deemed prohibited systems, high-risk systems or low-risk systems. It is entirely possible that many systems, such as those which recommend which products you might wish to purchase on a shopping website or which song you might like to listen to on a music app, although falling within a definition of AI systems as set forth in the EU AI Act and thus being subject to the EU AI Act, may be systems suitable for designation as low risk AI systems within the categorisation process established by the EU AI Act and therefore subject to minimal regulation in practice.

The EU appears to be moving forward apace with a new Act to regulate AI. The European Commission has proposed a full

draft of an EU AI Act. The EU Council has considered the European Commission’s proposal and issued its own text. The European Parliament has now also finalised its own wording of the EU AI Act. The European Commission, the EU Council and the European Parliament have now entered into discussions on a proposed compromise text on which they all can agree. Undoubtedly in the context of these discussions, the three EU bodies will take into careful consideration the OECD’s updated 2023 definition of AI system issued by the OECD subsequent to their own proposals.

Throughout the genesis of the EU AI Act to date, the definition of AI system, that is, the definition of that which is to be the subject of the regulation, has been totally rewritten, indeed replaced, at each fundamental stage of the process. This is perhaps not surprising given the fast-moving nature of the technology at issue as well as the potentially conflicting EU goals of protecting fundamental rights, fostering economic development and innovation and creating a law that is as future proof as possible. In order to achieve these goals in a harmonised manner to the greatest extent possible a definition of AI system that is clear and cannot be modified outside the EU legislative process, a definition which is sufficiently broad to accommodate future technological developments and a definition that focuses on systems that make predictions, recommendations and decisions would appear to be both the most practical and the most appropriate. Those concerned that such a definition of AI system, or the more expansive new 2023 definition of the OECD referring to outputs ‘such as predictions, content, recommendations, or decisions’, would be overly broad and thereby stifle innovation and economic development can take at least some consolation in the fact that many software systems, although falling within the definition of AI system, and therefore subject to the EU AI Act, could well be

categorised as low-risk systems subject to minimal regulation.⁴⁸

If the EU does indeed adopt a broad definition of AI system in the final version of the EU AI Act the related critical question as to which AI systems fall into the low-risk category, which AI systems fall into the more highly regulated high-risk category and which AI systems will be entirely prohibited will remain. Other than the question of the definition of AI systems discussed in this paper, the question of the categorisation of whether a given AI system falls within the low-risk, high-risk or prohibited category is likely to be an equally complex issue to resolve.

References and notes

1. Turing, A. M. (1950) 'Computing Machinery and Intelligence', *Mind*, Vol. LIX, No. 236, pp. 433–60. It was in this article that Turing set forth and analysed what he called 'The Imitation Game', which is now often referred to as 'The Turing Test'.
2. McCarthy, J., Minsky, M., Rochester, N. and Shannon, C. E. (1955) 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence', available at <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf> (accessed 16th October, 2023).
3. For further information see, for example, Siebel, T. (2019) 'Digital Transformation: Survive and Thrive in an Era of Mass Extinction', RosettaBooks, New York.
4. Google Deepmind (n.d.) 'AlphaGo', available at <https://www.deepmind.com/research/highlighted-research/alphago> (accessed 16th October, 2023); Vincent, J. (27th November, 2019) 'Former Go Champion Beaten by DeepMind Retires after Declaring AI Invincible', *The Verge*, November 27, 2019, available at <https://www.theverge.com/2019/11/27/20985260/ai-go-alphago-lee-se-dol-retired-deepmind-defeat> (accessed 16th October, 2023).
5. Google Deepmind (18th October, 2017) 'AlphaGo Zero: Starting from Scratch', available at <https://www.deepmind.com/blog/alphago-zero-starting-from-scratch> (accessed 16th October, 2023).
6. See, for example, Mollick, E. (14th December, 2022) 'ChatGPT Is a Tipping Point for AI', *Harvard Business Review*, available at <https://hbr.org/2022/12/chatgpt-is-a-tipping-point-for-ai> (accessed 16th October, 2023).
7. European Commission Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, 21st April, 2021. The proposed EU regulation governing the use of AI in its various draft forms discussed in this paper is hereinafter referred to generally as the 'EU AI Act'.
8. See for example Report of the Committee on Legal Affairs, European Parliament, with recommendations to the Commission on Civil Law Rules on Robotics, 27th January, 2017; European Commission Independent Report by a High Level Expert Group on Ethics Guidelines for Trustworthy AI, 8th April, 2019; and European Commission White Paper On Artificial Intelligence — A European Approach to Excellence and Trust, 19th February, 2022.
9. European Commission Proposal, ref 7 above, Section 1.1, Explanatory Memorandum.
10. *Ibid.*, Section 5.2.1, Explanatory Memorandum.
11. Although there is no definitive or exhaustive list of fundamental EU values, the Treaty on European Union states: 'The Union is founded on the values of respect for human dignity, freedom, democracy, equality, the rule of law and respect for human rights, including the rights of persons belonging to minorities. These values are common to the Member States in a society in which pluralism, non-discrimination, tolerance, justice, solidarity and equality between women and men prevail'. (Article 2 of the Treaty on European Union, available at https://eur-lex.europa.eu/resource.html?uri=cellar:2bf140bf-a3f8-4ab2-b506-fd71826e6da6.0023.02/DOC_1&format=PDF [accessed 16th October, 2023]). The Charter of Fundamental Rights of the European Union (available at https://www.europarl.europa.eu/charter/pdf/text_en.pdf [accessed 16th October, 2023]) also sets out a range of civil, political, economic, social and cultural rights and principles that the EU and its member states must respect and promote when implementing EU law.
12. European Commission Proposal, ref 7 above, draft EU AI Act, Article 5, Title II.
13. *Ibid.*, Section 5.2.1, Explanatory Memorandum. Note the distinction between prohibited systems 'contravening' EU values and high-risk systems 'posing a significant risk' to the fundamental rights of persons.
14. *Ibid.*, draft EU AI Act, Article 71, Section 3, Title X.
15. *Ibid.*, draft EU AI Act, Article 2(1), Title I.
16. *Ibid.*, draft EU AI Act, Article 3(1), Title I.
17. *Ibid.*, draft EU AI Act, Annex I.
18. *Ibid.*, Explanatory Memorandum, Section 5.2.1.
19. *Ibid.*, Article 3(1).
20. *Ibid.*
21. *Ibid.*
22. *Ibid.*, Preamble (6).
23. *Ibid.*
24. Council of the European Union. Interinstitutional File: 2021/0106(COD), 25th November, 2022. In the EU Council, the examination of the proposal was carried out in the EU's Working Party on Telecommunications and Information Society (WP TELECOM) which started discussing the proposal under the Portuguese Presidency, with meetings and workshops held between April and June 2021. The work on the proposal continued under the Slovenian Presidency, which drafted the first, partial compromise proposal covering Articles 1–7 and Annexes I–III. The Slovenian Presidency organised a half a day informal EU Council of telecom ministers devoted exclusively

- to the EU AI Act proposal during which the ministers confirmed their support for the horizontal and human-centric approach to regulating AI. The French Presidency continued the examination process, and by the end of its term, it redrafted the remaining parts of the text and presented the entire first consolidated compromise proposal on the EU AI Act on 17th June 2022. On 5th July, 2022, the Czech Presidency held a policy debate in WP TELECOM on the basis of a policy options paper, the outcomes of which were used to prepare the second compromise text. Based on the reactions of the delegations to this compromise, the Czech Presidency prepared the third compromise text, which was presented and discussed in WP TELECOM on 22nd and 29th September, 2022. After these discussions, the delegations were asked to provide further written comments, which were used by the Czech Presidency to draft the fourth compromise proposal. Based on the discussions of the fourth compromise proposal held in WP TELECOM on 25th October, 2022 and 8th November, 2022 and taking into account final written remarks from the EU member states, the Czech Presidency prepared the final version of the compromise text (Section III State of Play in the Council, Sections 1 and 2, EU Council Proposal).
25. *Ibid.*, draft EU AI Act, Article 3(1).
 26. *Ibid.*, Section IV Main Elements of The Compromise Proposal.
 27. *Ibid.*, Recital 6.
 28. *Ibid.*
 29. *Ibid.*
 30. *Ibid.*
 31. European Commission Proposal, ref 7 above, draft EU AI Act, Article 3(1).
 32. EU Council Proposal, ref 24 above, Recital 6(a).
 33. *Ibid.*
 34. Satariano, A. and Metz, C. (5th March, 2023, updated 6th March, 2023) 'Using AI to Detect Breast Cancer that Doctors Miss', *The New York Times*, available at <https://www.nytimes.com/2023/03/05/technology/artificial-intelligence-breast-cancer-detection.html> (accessed 16th October, 2023).
 35. Merriam-Webster (n.d.) 'infer', available at <https://www.merriam-webster.com/dictionary/infer> (accessed 15th March, 2023).
 36. Draft Compromise Amendments on the Draft Report; Proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts; COM(2021)0206 – C9 0146/2021–2021/0106(COD), Rapporteurs: Brando Benifei & Ioan-Dragoş Tudorache; Joint committee procedure – Rule 58 of the Rules of Procedure, 9th May, 2023.
 37. Amendments adopted by the European Parliament on 14th June, 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts.
 38. *Ibid.*, p. 136.
 39. As further background, see Bertuzzi, L. (7th March, 2023, updated 9th March, 2023) 'EU lawmakers set to settle on OECD definition for Artificial Intelligence', EURACTIV, available at <https://www.euractiv.com/section/artificial-intelligence/news/eu-lawmakers-set-to-settle-on-oecd-definition-for-artificial-intelligence/> (accessed 16th October, 2023).
 40. OECD Recommendation of the Council on Artificial Intelligence OECD Legal/0449 adopted on 22nd May, 2019, OECD Legal Instruments; Artificial Intelligence Risk Management Framework (AI RMF 1.0) (the 'OECD 2019 AI Recommendation'). The OECD AI Recommendation states that it provides 'the first intergovernmental standard for AI policies and a foundation on which to conduct further analysis and develop tools to support governments in their implementation efforts'.
 41. 15 U.S.C. 9401(3).
 42. National Institute of Standards and Technology, US Department of Commerce, (January, 2023) 'Artificial Intelligence Risk Management Framework (AI RMF 1.0)', p. 1, available at <https://doi.org/10.6028/NIST.AI.100-1> (accessed 16th October, 2023). The US National Institute of Standard's Risk Management Framework provides that it is to be a document which 'is designed to equip organizations and individuals [...] with approaches that increase the trustworthiness of AI systems, and to help foster the responsible design, development, deployment, and use of AI systems over time', p. 2.
 43. White House Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, October 30, 2023 (the 'White House AI Executive Order'). Neither the OECD AI Recommendation nor the US National Institute of Standards and Technology's AI publication have the force of law. US National Artificial Intelligence Initiative Act of 2020 has the force of law. The White House AI Executive Order, although binding, is, like all White House Executive Orders, subject to court challenge, legislative override and abrogation by a new administration.
 44. OECD AI Principles overview, available at <https://oecd.ai/en/ai-principles> (accessed 11th November, 2023).
 45. EU Commission Proposal, draft EU AI Act, Article 3(1), Title I (emphasis added).
 46. EU Council Proposal, draft EU AI Act, Article 3(1) (emphasis added).
 47. EU Parliament Proposal, see ref 37, p. 136.
 48. OECD AI Principles overview, see ref 44 above (emphasis added).