ELI
EUROPEAN
LAW
INSTITUTE

Guidelines on the Application of the Definition of an AI System in the AI Act: ELI Proposal for a Three-Factor Approach

Response of the ELI to the EU Commission's Consultation

EUROPEAN
LAW
INSTITUTE

The concept of 'AI system' under the new AI Act: Arguing for a Three-Factor Approach

The European Law Institute

The European Law Institute (ELI) is an independent non-profit organisation established to initiate, conduct and facilitate research, make recommendations and provide practical guidance in the field of European legal development. Building on the wealth of diverse legal traditions, its mission is the quest for better law-making in Europe and the enhancement of European legal integration. By its endeavours, ELI seeks to contribute to the formation of a more vigorous European legal community, integrating the achievements of the various legal cultures, endorsing the value of comparative knowledge, and taking a genuinely pan-European perspective. As such, its work covers all branches of the law: substantive and procedural; private and public.

ELI is committed to the principles of comprehensiveness and collaborative working, thus striving to bridge the oft-perceived gap between the different legal cultures, between public and private law, as well as between scholarship and practice. To further that commitment, it seeks to involve a diverse range of personalities, reflecting the richness of the legal traditions, legal disciplines and vocational frameworks found throughout Europe. ELI is also open to the use of different methodological approaches and to canvassing insights and perspectives from as wide an audience as possible of those who share its vision.

European Law Institute
Schottenring 16/ 175
1010 Vienna
Austria
Tel: + 43 1 4277 22101
E-mail: secretariat@europeanlawinstitute.eu
Website: www.europeanlawinstitute.eu

Approved by the ELI Council on 11.01.2024.
Final draft published on 11.01.2024.

ISBN: 978-3-9505495-3-9

Acknowledgement is due to the University of Vienna, which has generously hosted the ELI Secretariat under successive Framework Cooperation Agreements since 2011.

universität
wien

Table of Contents

Acknowledgements

Team

Authors

Bernhard Nessler (Research Manager Intelligent Systems and Certification of AI at the Software Competence Center Hagenberg, lecturer at the Institute for Machine Learning at the Johannes Kepler University Linz and Vice-President of the Austrian Society for Artificial Intelligence (ASAI), Austria)

Christiane Wendehorst (Professor of Civil Law and Deputy Head of the Department for Innovation and Digitalisation in Law at the Faculty of Law of the University of Vienna, Austria)

Advisory Committee

Marc Clément

Piotr Machnikowski

Pascal Pichonnaz

Teresa Rodríguez de las Heras Ballell

Rania Wazir

Project Assistants

Alexander Aufreiter (Legal Researcher Intelligent Systems and Certification of AI at the Software Competence Center Hagenberg and Senior Consultant at ICON Wirtschaftstreuhand)

Gregor Aichinger (Research Associate at the LIT Law Lab of the Johannes Kepler University Linz and Legal Researcher Intelligent Systems and Certification of AI at the Software Competence Center Hagenberg)

ELI Project Officer

Marta Lages de Almeida

Executive Summary

The AI Act represents a significant milestone in the European Union's efforts to regulate artificial intelligence (AI), with its primary goal being to establish a legal framework for trustworthy, human-centred AI. Central to the AI Act's effectiveness is the definition of 'AI system', a concept that has sparked considerable debate. In the final version of the AI Act, the European legislator has opted for a concept of 'AI system' that closely aligns with the revised Organisation for Economic Cooperation and Development (OECD) definition from November 2023. This alignment suggests reciprocal influence between the discussions surrounding the AI Act and the OECD's work. While international consistency in terminology is beneficial, the definition's vagueness and lack of clarity are surprising.

This consultation response critically examines the definition's technical and legal ambiguities and proposes a systematic Three-Factor Approach to delineate AI systems for the purpose of the AI Act. The approach evaluates systems based on three key dimensions:

Factor I — **The amount of data or domain-specific knowledge that went into development:** the data or domain-specific expert knowledge that went into the system's development, ranging from data-driven models to those leveraging expert rule-based programming.

Factor II — **The extent to which know-how is created during operation:** the presence of goal-oriented optimisation or search algorithms in the operation of the system, differentiating simple forward computations from systems capable of generating new knowledge on 'how to' solve a problem.

Factor III — **The degree of formal indeterminacy of outputs:** the formal indeterminacy in the system's outputs, where the system handles tasks that would, if performed by humans, involve discretion and require subjective judgement or creative interpretation.

The interpretation in the three factors acknowledges that the brief definition in the AI Act is intended to be an abstract picture that does not not dive into the technical details of current AI system implementations. The factors take the single elements of the definition and interpret them in the light of the current state-of-the-art in AI. In order to reflect a delineation of AI that is consistent with the overall goal of the AI Act and its Recitals, it transpires that the factors are interconnected, and a strong presence in one area can compensate for weaknesses in another. This flexible system balances technical neutrality and practical applicability, providing a more nuanced basis for categorising AI systems under the AI Act.

An AI system within the meaning of Art 3 No 1 AI Act is a machine-based system that …

Factor I:	Use of data or domain-specific expert knowledge in development
… infers, from the input it receives, how to generate outputs [input is training data]	
… for … implicit objectives [model development objectives]	
++	*data-driven programming (ML), expert-systems, ontologies, …*
+	*heuristic rules from experts*
0	*no domain–specific data analysis* *no expert knowledge*

Factor II:	Creation of new know-how during operation
… infers, from the input it receives, how to generate outputs [input is live data]	
… for explicit … objectives [system execution objectives]	
… may exhibit adaptiveness after deployment	
++	*goal oriented optimization, online reinforcement learning,* *inverse model calculations, deductive inference, search in complex data spaces*
+	*NP-complete algorithms, navigation, scheduling*
0	*pure forward calculations, deterministic or stochastic results, predefined calculation steps*

Factor III:	Formal indeterminacy of output
… designed to operate with varying levels of autonomy	
… outputs such as predictions, content, recommendations, or decisions	
… can influence physical or virtual environments	
++	*no formal criteria to define correctness of results, humans instead of the AI would exercise discretion, subjectivity in generation of outputs*
+	*variability of outputs but no subjective discretion; fixed distribution of results*
0	*clearly formally defined outcomes, variability of results are perceived as substantially identical by humans, general database systems, word processing programmes*

… cumulatively scores at least three "+" in the above Three-Factor-System.

I. Introduction

The recent entry into force of the European Union's AI Act has attracted significant international attention. Its aim is to establish a legal framework for trustworthy, human-centred AI. The AI Act's objectives are to ensure that AI serves human beings, remains under human control, benefits individuals and society, and does not infringe on fundamental rights or undermine the rule of law or democracy. However, the term 'AI' or 'AI system' remains one of the most ambiguous, both in legal discourse and in computer science, where a clear, universally accepted definition of AI is also lacking.[1]

The impact of the scope of the AI Act on the EU's status as a hub for business and innovation is still uncertain. A broader scope could deter companies concerned about additional regulatory burdens, potentially discouraging them from doing business in Europe. Paradoxically, however, a broader interpretation of 'AI system' could also make Europe more attractive for innovation: If compliance with the AI Act cannot be avoided by sidestepping certain emerging technologies, these technologies are more likely to be adopted. Considerations of technological neutrality[2] and the effectiveness of the AI Act also support a broader understanding of the definition, as the legal and ethical challenges posed by AI apply equally to all advanced algorithmic systems.[3]

[1] Cf Barr/Feigenbaum, The Handbook of Artificial Intelligence, 1981, p 3 f on the term itself, which is generally recognised in computer science but not on the delimitation; for further details see Winter et al, Trusted Artificial Intelligence - Towards Certification of Machine Learning Applications, 2022, p 11 ff, https://www.tuv.at/wp-content/uploads/2022/03/Whitepaper_Trusted-AI_TUeV-AUSTRIA_JKU.pdf.

[2] See In Detail Uncitral, Explanatory Note By The Uncitral Secretariat On The United Nations Convention On The Use Of Electronic Communications In International Contracts, 2007, P 26 F; Most Recently Uncitral, Taxonomy Of Legal Issues Related To The Digital Economy, 2023.

[3] Data Ethics Commission of the Federal Government, Report of the Data Ethics Commission, 2019, p 24 ff, 159 ff.

II. History of the definition in Article 3 No 1 AI Act

From the start of the legislative procedure, the definition of 'AI system' in the AI Act was one of the most controversial points. The European Commission's **initial proposal from April 2021** defined an AI system as:

> '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'.

Annex I, which could be modified by the Commission through delegated acts, covered three major groups of technologies and concepts:

> '(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; and (c) statistical approaches, Bayesian estimation, search and optimization methods.'

During the legislative process, there were efforts to limit the scope to machine learning, arguing that only data-driven learning (as opposed to rule-based programming) introduces the level of complexity, opacity and limited predictability that warrants new regulation.[4] Conversely the Commission's definition was criticised for deviating significantly from the OECD's more technology-neutral definition.[5] The **2019 OECD AI recommendations** stated the following definition:

> 'An AI system 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. AI systems are designed to operate with varying levels of autonomy.'[6]

Reflecting this perspective, **the European Parliament's position** adopted in June 2023[7] defined an AI system, without any restriction to specific technologies, as:

> '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'.

In November 2023, however, the OECD countries agreed on a **revised OECD definition**:

> '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

---

[4] See KI-Bundesverband, Feedback to the European Commission's regulation proposal on the Artificial Intelligence Act, 6 August 2021; Bitkom, Bitkom principles for the Artificial Intelligence (AI) Act, 4 August 2021; DIHK, Stellungnahme Deutscher Industrie- und Handelskammertag, 5 August 2021; see also Steege MMr 2022, 926 in favour of greater restriction; see also Wiebe BB 2022, 899; Spindler CR 2021, 361 (373). AA for example VZBV, Artificial Intelligence needs real world regulation, 5 August 2021.

[5] Digitaleurope, AI Act trilogues: A vision for futureproofing, governance and innovation in Europe, 16.10.2023, p 1, 5; ZVEI, Statement on the EU Commission's proposal for a Regulation laying down harmonised rules on Artificial Intelligence ("AI Act"), 8.2021, p 3; see also Becker/Feuerstack MMR 2024, 22 (23).

[6] OECD, Recommendation of the Council on Artificial Intelligence, 2019, https://www.oecd.org/berlin/presse/Empfehlung-des-Rats-zu-kuenstlicher-Intelligenz.pdf.

[7] Parliament doc A9-0188/2023.

> as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment'.[8]

During the trilogue, the parliamentary definition was adapted to the revised OECD definition. As a result, an 'artificial intelligence system' (AI system) is now defined in **Article 3 No 1 AI Act** as:

> 'a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments'.

Recital 12 AI Act provides additional context with regard to this definition.

[8] OECD, Explanatory memorandum on the updated OECD definition of an AI system, 2024.

III. Elements of the definition

1. AI as 'a machine-based system …'

a) Distinguishing AI from biological systems

The definition describes an AI system, first and foremost, as a 'machine-based system', avoiding terms such as 'software', 'algorithmic system' or 'IT system'. According to Recital 12, this wording indicates that AI systems run on machines. However, the exact implications remain unclear, especially since the term 'machine' is not further defined. The description appears to **exclude all biological systems**, even those that are artificially produced or manipulated (eg cell structures from brain organoids). Whether this exclusion will be particularly helpful given the rapid convergence between computer science and life sciences remains to be seen.

b) Distinguishing AI from other product components

AI systems can function as stand-alone products or as a component of other products. In the latter case, AI systems may be either physically integrated (**embedded**) or serve the product's functionality without being integrated (**non-embedded**). This raises the question of whether only the component fulfilling the special characteristics of an AI system should be classified as an 'AI system' within the meaning of the AI Act, or whether the entire product should be classified as such.[9]

There is no uniform answer to this question. Article 6(1) AI Act, which distinguishes between AI systems as products and AI systems as safety components of other products, suggests that most hardware components (eg the machine in which the AI system is embedded) are not considered 'AI systems'. Thus, under the AI Act, a robot is regarded as a machine with an embedded AI system. However, since the conformity of an AI system must always be assessed within its specific hardware and software environment, this distinction is not particularly significant in practice.[10]

For software products (eg medical diagnostic systems or software assessing the creditworthiness of individuals), it is often even more difficult to distinguish between system components that qualify as AI and those that do not. For instance, in the case of recruitment software, it may be difficult to separate AI-supported application screening and ranking from general application document management. However, in cases of functionally distinct modules – such as an AI recruitment module within a general document management system – other components do not become part of the AI system if they are connected to it via an open interface, allowing for a clear identification of the inputs and outputs of the AI system.

2. '… that is designed to operate with varying levels of autonomy …'

Machine-based systems are only classified as AI systems if they exhibit specific characteristics associated with AI. One such characteristic, mentioned in the definition in Article 3 No 1 of the AI Act, is that AI systems are designed to operate with varying levels of autonomy.

[9] Difficulties or questions regarding the exact delimitation can arise, for example, with a version of 'Spot' (an 'artificially intelligent' robot dog from Boston Dynamics), which can also be equipped with an LLM, https://www.youtube.com/watch?v=djzOBZUFzTw.

[10] See also Directive (EU) 2024/2853 of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC, which includes embedded software, including any AI, as well as self-standing software

a) Independence from human intervention

According to Recital 12, this means that AI systems can function, to some extent, **independently of human intervention** – similar to the automation process in a Turing machine. The technical frame of this autonomy is apparently limited only by the general limits of computability, as defined by Turing machines. This explanation in Recital 12 appears to suggest that any automated data processing carried out by deterministic or stochastic algorithms and any existing IT system could be classified as an AI system, and even systems with zero autonomy – eg within the meaning of point 5.13 of Standard ISO/IEC 22989:2022 – might fall within the definition.[11]

This broad interpretation would be consistent with the wording of Article 3 No 1 AI Act, as the term 'varying' could be read as including 'zero'. It is further supported by the (revised) OECD definition as the OECD's Explanatory Memorandum states: 'AI system autonomy (contained in both the original and the revised definition of an AI system) means the degree to which a system can learn or act without human involvement following the delegation of autonomy and process automation by humans. (…)'.

b) Nature of the outputs

On the other hand, it is difficult to accept that the definition in Article 3 No 1 AI Act simply equates 'autonomy' with 'automation'. A more convincing reading of this element of the definition might therefore be to interpret it as closely **linked to the nature of the outputs**, ie if the output is of a kind that was previously only produced by humans because it involves a high degree of 'discretion' when performed by humans without the support of AI, we consider this to be 'autonomy' (see below under III. 6.).

Note that the AI-system itself is a strict algorithmic system that does not have any freedom for discretion in its processing. Mere stochasticity (randomness of dices) should not be mistaken as intentional discretion. The room for discretion that a human would have in creating such output is 'eaten up' intentionally or unintentionally by the design process of the AI, which is controlled by human engineers.

3. '… and that may exhibit adaptiveness after deployment …'

Article 3 No 1 AI Act defines an AI system as a machine-based system which, inter alia, 'may exhibit adaptiveness after deployment'. To better understand how a system can exhibit adaptiveness after deployment, it is helpful to look at the legal interpretation of the meaning of the word and the technical background.

a) Meaning of 'adaptiveness'

According to Recital 12, 'the adaptiveness that an AI system could exhibit after deployment refers to the self-learning capabilities, allowing the system to change while in use', ie **'adaptiveness' is understood as 'model optimisation during operation'**. By way of comparison, the Explanatory Memorandum to the (revised) OECD definition is somewhat ambiguous in this regard as it appears to blend the development phase and the deployment phase. It states: 'Adaptiveness (contained in the revised definition of an AI system) is usually related to AI systems based on machine learning that can continue to evolve after initial development. The system modifies its behaviour through direct interaction with input and data before or after deployment. ... AI systems can be trained once, periodically, or continually (...)'.[12]

In a technical sense, a system may be called adaptive if it changes in any way over time by factual mechanisms regardless of whether or not that adaptation serves any apparent goal. In the context of an AI-system which iterates the process of inferring outputs from inputs according to a certain deterministic or stochastic input-output-function, adaptation refers to a **change of that function as a side-effect of the iterated application**. This seemingly clear technical

[11] See also the levels of autonomy described in ISO/IEC 22989:2022, 5.13.
[12] OECD, Explanatory memorandum on the updated OECD definition of an AI system, 2024, p 6 ff.

definition, however, becomes blurred when it comes to a multi-step processing of sequential inputs, as, for example, in an interactive session of a typical chatbot. If every single question-answer pair were seen as an input-output frame, then obviously the system would be adaptive during the conversation, as the past messages of the conversation are stored in the so-called context which gives the chatbot the ability to take earlier messages from the context into account. If, however, a whole conversation is seen as a multi-step input-output frame, then the system is not adaptive, as its state is again the same at the beginning of each new conversation. The latter perspective is the current common understanding in the AI community. As every technical system can typically be reset to its initial state at any desired point in time, like a chatbot at the beginning of each conversation, it is largely a question of the perspective of the system if a system is to be seen as adaptive, or if the single frames of the input-output function stretch over the time period between two state resets.

In view of the very broad technical understanding of adaptiveness, it likely makes sense to interpret the meaning in the context of the AI Act **as goal-oriented adaptation with respect to a specific objective**. From an external point of view, it is technically impossible to decide whether an adaptation made by a system is goal-oriented or not (see eg evolution). However, unlike biology, technical systems, including AI systems, are designed by human engineers according to typical design patterns. A technical introspection of the source code of an AI system will, therefore, usually reveal whether the internal model is indeed changed from one input-output frame to the next in a way that reflects the intention of the system's designer to improve the system with respect to a certain goal. In clear cases, these goals are explicitly encoded.

It should also be noted that oftentimes, the attribute 'adaptive' is merely used as a selling argument for any kind of system, even if the core of the system, ie the algorithmic model of the function, remains the same and the appearance of adaptation merely arises from the explicit memorisation of values or facts. This should not be understood as 'adaptiveness' in the sense of the definition in the AI Act.

b) Adaptiveness in machine learning

Adaptiveness is a feature that is often mentioned in the context of machine learning. In the pre-deployment stage of a machine learning (ML) based AI application,[13] the focus is on training the ML model. The model's parameters are carefully adjusted and optimised using statistical inference based on a specific training data set and a formally explicit goal. Complex optimisation methods, typically monitored and controlled by humans, are employed for this purpose. Once training is complete and the system is deployed, there are three possible scenarios: the use of fixed models, fixed models with updates, or online[14] self-optimisation of the AI system.

When using **fixed models**, the fully trained ML model[15] is integrated into an AI application, which typically includes other components, such as API access points and graphical user interfaces (GUI), alongside the fixed ML model. In this case, the model is no longer trained after deployment and the algorithms used for training are not part of the final AI system. During operation, the system applies the fixed pre-trained model to inputs, generating outputs such as decisions, predictions, content or recommendations based on what it learned during development.

When using **fixed models with updates**, the trained model is further developed after the initial version has been fixed and deployed. Additional data – such as data collected during the system's use - can be employed by human ML-engineers to develop the model further, by including this additional data in subsequent training and optimisation procedures. Even the model's architecture could be modified

[13] The term 'AI application' is deliberately used in the context of this consultation response to refer to software and systems that are commonly referred to as 'AI' and is not congruent with the term 'AI system' within the meaning of Art 3 No 1 of the AI Regulation.

[14] In computer science, the term 'online' means 'in operation'.

[15] In this state, the model is often referred to as 'frozen', although there is no actual freezing step.

and partially or fully retrained from scratch. The improved version is then checked and tested by humans and eventually deployed, either manually or automatically, to the instances of the AI system already in use, much like automatic software updates that are now commonplace. Even if these updates to the fixed models are automated and occur frequently, this process, which is monitored and controlled by developers, should not be mistaken for the (online) adaptiveness of the deployed AI system. Regular updates may give non-technical users the impression that the model has 'learned' during use. However, it is actually a 'newer', subsequent model that has been trained and updated based on the new data that was acquired from the recent interaction of the AI system with all or some of its users.[16]

Currently, there are only a few fields of application where ML systems genuinely allow or require (certain) **online self-optimisations during operation** based on real-time inputs and outputs. Notable examples include very restricted and simple cases of reinforcement learning methods (eg for recommendation systems) and adaptive control systems in robotics and electrical engineering (eg Kalman filter). The usage of such self-optimisation in hitherto real-world AI systems is greatly limited to very narrow low-dimensional adaptations. A high-dimensional ML optimisation processes (learning) always needs to be guided, analysed and tested by humans as the quality and results of the optimisation process is, in general, still impossible to predict (alignment problem) according to the current scientific state-of-the-art. AI systems based on current Large Language Models or other GPAI-ML-models are typically not self-optimising.

c) Adaptiveness as an optional feature of AI?

The various language versions introduce some ambiguity as to whether adaptiveness, thus understood, is a mandatory feature for an AI system, or whether it is simply an optional characteristic. The fact that adaptiveness is mentioned as a separate element of the definition in Article 3 No 1 of the AI Act might be read as suggesting that post-implementation adaptiveness is of some importance for a system to qualify as an AI system. Ultimately, however, the more convincing reading is to conclude that **'may' indicates optionality** rather than a necessary requirement.

If adaptiveness after deployment were mandatory, the AI Act's scope would be significantly narrowed with regard to ML-based AI systems, as most ML models currently on the market are clearly non-adaptive. Once they are deployed and in use, most ML-based AI systems lack the ability to learn from data or adapt their parameters, because the optimisation methods required for this are not present in these AI systems, as explained earlier (see above b).[17] This strongly suggests that the mention of adaptiveness is intended to be illustrative, clarifying that such systems can exist, rather than being a strict requirement.

Although the feature of adaptiveness, understood in this way, is optional, the presence of adaptation in the narrow understanding of an **explicit goal-oriented online-optimisation** could be seen as a **strong indicator in favour of a system's qualification as an AI system**.

[16] See also Nessler/Aufreiter/Aichinger, In Scope? - The Definition of 'AI System' in the AI Act, in: The First Austrian Symposium on AI, Robotics, and Vision (AIROV24) (forthcoming), p 4.

[17] For further information on the 'learning' of LLM, see Mayrhofer/Nessler/Bieber/Fister/Homar/Tumpel, ChatGPT, Gemini & Co - Große Sprachmodelle und Recht/ Nessler/Schmid/Mederitsch/Aichinger, 2024, p 15 (p 40).

4. '… and that, for explicit or implicit objectives …'

The reference to explicit or implicit objectives in the context of the AI system's capability to infer is another element of the definition whose role is not immediately obvious.

a) Human objectives

Any algorithmic operation, ie any software system or functional component thereof, may be interpreted as fulfilling certain explicit or implicit goals or objectives, understood as the wider goals or **objectives pursued by the parties developing or deploying the system**. Whether such objectives are 'explicit' or 'implicit' is usually a matter of (human) perspective, as the system itself does not share these objectives. For example, multiplying an input by six and subsequently dividing the result by five may be said to pursue the implicit goal of computing the sales price including a VAT rate of 20%, yet the algorithmic process is not aware of the meaning of taxes. The OECD's Explanatory Memorandum sets out that an explicit human-defined objective would be a game-playing system designed to win a game, while an implicit objective could be an instruction to stop at a red traffic light. In the latter case, the deeper, underlying objective that the system developers had in mind – complying with the law and avoiding accidents – is not explicitly described in the algorithm.[18]

b) Algorithmic interpretation of objectives

However, the reference to 'objectives' could also be understood as referring to groups of algorithmic methods that are explicitly intended to pursue a certain predefined objective in the sense that a certain argument value is sought such that said objective is fulfilled or approximated. These are goal searching algorithms, optimisation algorithms and so-called solvers. From a technological perspective, it makes sense to differentiate between **not-objective-oriented forward calculations**, where the result of a model function is simply computed using the sequence of operations predefined in the model,[19] and **objective-oriented backward calculations**, which search for an argument where the model function meets certain predefined goals.[20]

Explicit goal value search is commonly used in the **execution of logic- and knowledge-based systems**. By contrast, in AI applications that generate outputs using fixed ML models, no target value search occurs in the execution phase. The fixed model is evaluated at the input point through a simple forward calculation. As noted below in the discussion of inference (see III. 5), target value search is employed in certain recommender systems, adaptive rule-based expert systems, and some developments built on top of large language models (LLMs). Therefore, the use of objective-pursuing algorithms in the deployed AI system could serve as a further criterion in favour of a classification as an AI system.

This does not mean that 'objectives' do not play any role at all in the context of ML. Typical ML training algorithms are optimisation algorithms, like gradient descent, or hypothesis search. Those **optimisation algorithms are used for the development of ML systems**. The objectives are carefully designed by ML engineers for the purpose of extracting useful statistical knowledge from the training data. Oftentimes, the engineering of ML models comprises training stages with different objectives that are applied consecutively. The development of an LLM typically comprises a so-called pre-training phase, and one or multiple fine-tuning phases. In the pre-training phase, a typical objective is the quality of the next-word-prediction, whereas the fine-tuning phase might target question-answering behaviour or tune the style of the generated text. It should be noted that all those training objectives from the development phase, while explicit in the training process, are

[18] OECD, Explanatory Memorandum, on the updated OECD definition of an AI system, 2024, p 6 ff p 7.
[19] Given x, calculate y = f(x).
[20] Wanted x, so that f(x) = 0, or wanted x, so that f(x) is maximum.

not part of the AI system in its deployed form. Yet these training objectives have shaped the finally deployed model of the AI system and are therefore **implicitly contained in the final AI-system**.

c) The role of objectives in the definition of AI

If the reference to explicit or implicit objectives is understood as referring to the objectives pursued by the parties developing or deploying the system, this element of the definition fails to fulfil any delimiting function as there will always be some objective that is pursued and nobody would develop or deploy a system without any – explicit or implicit – objective in mind.

Recital 12 states 'that AI systems can operate according to explicit defined objectives or to implicit objectives' and goes on to explain that the objectives of the AI system may differ from its intended purpose in a specific context. This seems to indicate that the objectives referred to could be the algorithmic objectives as described above, ie either **explicit objective-oriented backward calculations** (in the deployment phase) or **implicit objectives** no longer visible in the trained model as they were only visible as training objectives in the development phase. This reading of the definition supports and strengthens the delimiting function of the 'how to' phrase.

5. '… infers, from the input it receives, how to generate outputs …'

The final version of the definition also requires an AI system to infer, 'from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions'. Recital 12 emphasises that a 'key characteristic of AI systems is their capability to infer' and explains that the techniques enabling such inferences include ML approaches that learn from data how to achieve certain objectives, as well as logic- and knowledge-based approaches that infer from coded knowledge or symbolic representations of tasks. Recital 12 notes that the concept of an 'AI system' should not cover systems 'based on the rules defined solely by natural persons to automatically execute operations'.

a) Two possible interpretations

According to the wording of the final definition, an AI system does not have to infer the outputs themselves, but rather **'how' to generate the outputs**. Since Article 3 No 1 only mentions the capability to infer once, there are two possible interpretations:

The first, literal interpretation of Article 3 No 1 suggests that the defining characteristic of an AI system is its 'ability to derive models and algorithms'. Following this interpretation, most deep learning applications commonly referred to as AI, which lack the ability to derive models after deployment (see III. 3.), would not be considered AI systems within the meaning of Article 3 No 1 AI Act. Only the few ML systems that retain the automated ability to derive models and algorithms during operation would qualify. Examples of such systems would include, recommender systems based on online reinforcement learning, adaptive control systems in robotics, and the latest developments in language models, such as AutoGPT, Chain-of-Thought or Tree-of-Thought, where the system independently finds creative ways to fulfil user-defined goals. Although this interpretation aligns with the precise wording of the AI Act, it would significantly limit its scope.

The second interpretation largely ignores the word 'how', suggesting that 'infer' relates directly to generating outputs. This would imply that, to fall under the definition, an AI system only needs the 'ability to generate outputs'. Since almost all IT systems can generate outputs, this interpretation would vastly expand the scope of the AI Act.

b) Development and deployment phase

Again, a solution to this interpretive issue likely lies in the Explanatory Memorandum to the OECD's AI definition. Regarding the meaning of 'inference', the OECD explains: 'The concept of "inference" generally refers to the step in which a system generates an output from its inputs, typically after deployment… When performed during the build phase, inference, in this sense, is often used to evaluate a version of a model, particularly in the machine learning context. In the context of this explanatory memorandum, "infer how to generate outputs" should be understood as also referring to the build phase of the AI system, in which a model is derived from inputs/data.'[21]

There is substantial evidence to suggest that the statements in the definition refer to entirely different phases or different aspects. For **ML approaches**, 'inference' could relate to the development process ('build phase') and the **inference of models and algorithms from training data**. Recital 12 initially emphasises that inference refers to 'the process of obtaining the outputs (...) and to a capability of AI systems to derive models or algorithms (…) from inputs or data'. It then explains that the relevant techniques that 'enable inference while building an AI system include machine learning approaches that learn from data how to achieve certain objectives'. In this context, the term 'inputs' likely refers not to 'input data', as defined in Article 3 No 33 AI Act, but rather to training data, as defined in Article 3 No 29 AI Act, or functionally similar inputs – unless, of course, the AI Act only applies to systems that exhibit adaptiveness in the strictest sense after deployment (see point 3 above).

The use of the term 'infer' appears to shift when the AI Act mentions **logic-based and knowledge-based approaches**, where the focus moves to the operation phase and the **inference of outputs**. Such systems typically consist of a knowledge base and an inference engine, which produces results by drawing conclusions based on the knowledge base. The knowledge base, usually created by human experts, represents entities and logical relationships relevant to the application problem. These entities and relationships are formalised, based on rules, ontologies, or knowledge graphs. Once developed and operational, the inference engine derives solutions from coded knowledge or symbolic representation of the task at hand. For logic- and knowledge-based systems, the 'inputs' are indeed the input data as defined in Article 3 No 33 AI Act.

6. '… such as predictions, content, recommendations or decisions …'

An AI system is defined as a system capable of generating output such as predictions, content, recommendations or decisions.

The term '**predictions**' refers to statements about the likelihood (probability) of future events occurring or not occurring, making them primarily forward-looking. '**Content**' refers to generative AI systems that produce texts, images, videos, music, code or other forms of content. '**Recommendations**' involve formulating or selecting a course of action aimed at a specific goal (eg patient recovery) or multiple goals. In contrast, 'decisions' are not directed at a separate actor (eg a human being operating the AI system) to suggest action. Instead, decisions involve executing the action itself or at least determining it to such an extent that the executing actor has no discretion.[22] The term 'decisions' should be interpreted broadly, covering a wide range of legal (eg a decision to reject a contract) or factual actions (eg a robot's decision to move left). The different types of outputs often overlap and the list is by no means exhaustive.

Generating predictions, content, recommendations or (complex) decisions typically involves activities where a person (without support of AI) exercises some degree of discretion, unlike tasks that are purely computational. Since the desired output is somewhat undefined, different individuals might interpret and solve these tasks differently. Therefore, a system's ability to solve tasks **without a formal definition of the solution** – where a human would normally exercise discretion – could be a key criterion for classifying it as an AI system.

[21] OECD, Explanatory memorandum on the updated OECD definition of an AI system, 2024, p 9.
[22] OECD, Explanatory memorandum on the updated OECD definition of an AI system, 2024, p 9.

7. '… that can influence physical or virtual environments.'

Finally, the output of an AI system must be capable of influencing physical or virtual environments. According to Recital 12, 'environments' refers to 'the contexts in which AI systems operate, whereas outputs reflect different functions performed by AI systems'. A physical environment is influenced when the AI system's output, through actuators, directly triggers a mechanical reaction (eg in robotics). A virtual environment is influenced when the output is used as input for another algorithmic system. Expert systems and similar technologies that provide recommendations to human users (eg via a display) influence human sensory perceptions and cognitive processes (eg the eyes and brain).

Overall, this characteristic appears more illustrative and it remains unclear how it could help in delineating AI systems from other systems in ambiguous cases.

IV. Critical evaluation and Three-Factor Approach

1. The difficulty of drawing a line between AI systems and other systems

As demonstrated, practically no single element of the definition in Article 3 No 1 of the AI Act is effective in clearly distinguishing AI systems from other systems. It is, however, essential to draw a line between AI systems and other systems.

Recital 12 states that: '… the definition should be based on key characteristics of AI systems that distinguish it from simpler traditional software systems or programming approaches and should not cover systems that are based on the rules defined solely by natural persons to automatically execute operations'.

It should be noted that the distinction between 'simpler' and 'more complicated' systems can be misleading, as ML techniques can be very simple, whereas other traditional programming approaches can be highly complex. The OECD Explanatory Memorandum rightly states that systems of any degree of complexity are covered.[23] Also, ML is not exactly a new technique, but has only seen an unprecedented boost with the exponential increase in computing power and available data. It is, therefore, hardly possible to draw a clear line between 'simpler' or 'traditional' approaches and AI systems. The Explanatory Memorandum to the revised OECD definition also appears to suggest a broad approach, stating that 'AI models can be built manually by human programmers or automatically through, for example, unsupervised, supervised, or reinforcement machine learning techniques'.

2. The Three-Factor Approach

This consultation response , with a view to the guidelines to be prepared and issued by the Commission, proposes a systematisation based on three distinct factors.[24] These factors are mentioned in the definition in Article 3 No 1 and in Recital 12. They relate to the development, to the functioning and to the areas of application of a machine-based system: (I.) the use of data or (functionally similar) domain-specific expert knowledge in development; (II.) the creation of new know-how through goal-oriented online optimisation during system operation; and (III.) the lack of formal definition in the system's outputs.

The intensity of each factor – also shown in the table below (under point 4) – is classified as follows: '++' indicates a strong association with AI, '+' a moderate association, and '0' indicates no association with AI.

a) Factor I – data or domain-specific expert knowledge in development

The first factor examines the amount of data or domain-specific expert knowledge that went into the development process, for example, the extent to which statistical analysis of (training) data sets was involved. This follows from Recital 12, which distinguishes AI systems from 'traditional software systems or programming approaches', emphasising that AI systems are built using techniques that 'learn from data how to achieve certain objectives' or that 'infer from encoded knowledge or symbolic representation of the task to be solved'. Factor I can be viewed along a spectrum:

---

[23] In contrast, OECD, Explanatory memorandum on the updated OECD definition of an AI system, 2024, p 9, where it is emphasised that systems of any degree of complexity are covered.

[24] See Wendehorst/Nessler/Aufreiter/Aichinger, Der Begriff des „KI-Systems" unter der neuen KI-VO - Vorschlag eines „Drei-Faktor-Ansatzes" zur Beseitigung von juristischen und technischen Ungereimtheiten, Zeitschrift für IT-Recht und Recht der Digitalisierung, Heft 7 2024, 605-614; Martini/Wendehorst, KI-VO/Wendehorst, 2024, KI-VO Art 1 para 50 ff.

At the upper end of the spectrum (I ++), representing primarily **data-driven programming**, are various machine learning methods using optimisation algorithms and extensive training data to create the system, even if this process is meticulously monitored and controlled by ML developers. This category also includes **logic- and knowledge-based approaches**, where domain-specific know-how, experience and judgement from experts are explicitly gathered, processed with automation methods and coded into the system (expert systems).

In the middle of the spectrum (I +) lies the manual programming of the system by natural persons, whereby the programmers implicitly **consult statistical analysis of historical data and/or rely on their expert knowledge** in a particular domain.

At the lower end of the spectrum (I 0) are programmes and systems created without specific know-how about the application domain and without any significant analysis of data, such as general database systems or word processing programmes.

When assessing the role of data or domain-specific knowledge during development, it is important to note that the use of transparent explicit data material, such as databases that are only searched and individually referenced at the time of application (eg telephone directories or digitised maps), does not increase the factor I rating. On the other hand, it is considered relevant for factor I if data from the ongoing operation of the application flows back into the development process and is used to further refine the model for subsequent updates (I ++). Whether the data are used only for dictionary look-ups (I 0) or whether (statistical) knowledge is extracted by processing the data (I ++) is usually very clear to distinguish, except perhaps for some very unusual cases of application

b) Factor II – creation of new know-how during operation

The second factor relates to the creation of new know-how during operation, which is closely connected with the differentiation between simple calculations of functional results (forward calculations) and complex optimisations to achieve certain objectives (backward calculations)[25] (see III. 4). This goal-oriented optimisation character may either be restricted to a single input-output computation or span across multiple inputs and future outputs. The latter could be called a goal-oriented adaption of the model itself (see III. 3). This factor also spans a spectrum.

At the upper end of this spectrum, representing goal-oriented optimisations (II ++), are systems that **generate new know-how** (ie the know-how as to 'how to generate outputs') during use and are **guided by abstract objectives**. Factor II closely aligns with the wording of Article 3 No 1 and Recital 12, encompassing deductive methods of symbolic AI and inductive methods of ML. These systems utilise optimisation algorithms and data-dependent solution heuristics, typically yielding approximations of a theoretical optimum or involve difficult trade-offs, often with ambiguous results. High-end examples of these systems include deductive expert systems, online reinforcement learning systems, adaptive control systems for robotics, online adaptive recommender systems, and any goal-oriented adaptation in response to generated outputs or online feedback. The results of these systems are heavily dependent on the heuristics used, the knowledge base or historical data.

In the mid-range (II +) are **optimisation systems,** such as navigation systems or target value searches in technical applications, as well as simple algorithmic solutions to typical NP-complete[26] problems like scheduling systems (eg for generating timetables or planning production processes with limited resources). The solution-finding process in these cases is computationally intensive but **formally well-defined**, typically not utilising additional historical data or excessive domain-specific heuristics.

[25] For example, the calculation of an inventory value at given valuation prices is a simple forward calculation. Conversely, determining valuation prices to achieve a specific inventory value, which can be ethically questionable, is a backward calculation.

[26] Complete in the class of non-deterministic polynomial time solvable problems.

At the lower end (II 0) are pure calculation algorithms that produce a deterministic result following a predefined sequence of calculation steps. As already explained (see III. 4), fixed deep learning systems, which are normally perceived as AI, are very simple calculation algorithms when it comes to their mode of operation. Classification systems, image recognition systems and even chatbots based on large language models (LLM) such as ChatGPT are such simple forward calculations and thus fall into this category (II 0) even if they sometimes incorporate random numbers. The complexity of the training of these models is accounted for in factor I (I ++).

c) Factor III – degree of formal indeterminacy of outputs

The need to introduce the third factor arises primarily from the wording of Article 3 No 1, which states that AI systems produce 'outputs such as predictions, content, recommendations or decisions'. As explained in III.6, those **outputs lack a formal definition** and typically allow for a certain degree of discretion. This factor should therefore capture the extent to which different humans in the place of the AI system would be expected to generate substantially different outputs, due to their different experiences and personal views of the world. In the absence of an objective formal definition of the system's function, people may perceive various outcomes as subjectively accurate in the light of their individual motivations or perspectives.

At the upper end of the spectrum of formal indeterminacy or subjectivity (III ++), **multiple results may each be perceived as correct by different users**, and no formal criteria exist to verify the correctness of a result, at least not at the time of application. This is typically the case with predictions (eg the probability of a customer defaulting on a loan), recommendations (eg a film tailored to the user's preferences), content (eg texts, images, music) or decisions where a human being would typically exercise some discretion (eg whether to grant a loan).

In the middle of the spectrum (III +), there are systems whose results display **some degree of indeterminacy or variability**, which also humans perceive as uncertain. However, experienced individuals generally recognise the same distribution of possible outcomes, regardless of personal worldviews. Thus, subjective discretion is generally not involved. Examples include deciphering handwritten text, routing or scheduling problems in view of multiple equivalent solutions, or predicting (guessing) well-defined random events, such as the outcome of a roulette spin or a dice roll, within a given probability distribution.

At the lower end of the spectrum (III 0) are tasks with clearly defined outcomes, where each input can produce only one correct result (eg calculating the VAT charge for given revenues and VAT rates) or one correct response (eg sending a confirmation of receipt for an electronic order). Even if systems display some variability in their (intermediate) outputs, they can be categorized as III 0 if the effect of those different possible outputs (see III. 7) is perceived as substantially identical by humans. Examples include a robotic arm that repetitively lifts weights along a predefined path. The path is always perceived as the same, even though the steering commands vary with each repetition of the task due to changing environmental conditions.

3. Relationship of the three factors to each other

The three factors are largely independent from each other. For example, simple tasks, such as calculating VAT charges, could theoretically be solved through ML by leveraging numerous predefined sample calculations. Conversely, complex predictions or assessments, such as determining creditworthiness, can be programmed only using a rule-based approach, where humans manually assign point values to various parameters (eg income, home address, or frequency of overdrafts). Adaptiveness can be achieved not only through ML, but also through simple 'if-then' statements. However, certain combinations are typical: for example, an ML system generated through data-driven methods may require only simple forward calculations but is generally used for more complex tasks that cannot be formally defined.

a) The argument for a flexible system

Given this context, the most convincing approach is to view the three factors as interacting within a flexible system, where a **strong manifestation of one factor can compensate for a weak or even absent manifestation of another**. ML approaches and logic- and knowledge-based concepts are mentioned in nearly all drafts of the AI Act's definition and accompanying Recital 12. This suggests that the first and second factors are mutually interchangeable: in ML systems, the first factor is strong (I ++), while the second is usually not pronounced at all (II 0); in logic- and knowledge-based systems, the first factor is typically moderate (I +), and the second is strong (II ++).

A particularly challenging question is whether a strong manifestation of the third factor (III ++), ie the lack of a formal definition of outputs, can compensate for the absence of the aforementioned technologies. The answer to this question is crucial for determining whether many important applications, such as assessing loan applicants' creditworthiness or evaluating student performance, fall within the scope of the AI Act. To ensure the definition remains as technology neutral as possible (see I.), it seems advisable to answer this question in the affirmative, provided that considerable domain-specific expertise or statistical sources have been incorporated into the programming. Thus, imperative systems programmed by human domain experts should be included if they perform tasks with largely indeterminate outcomes, such as predictions, recommendations, or text generation, where humans typically exercise significant discretion.

b) Total score of three '+'

The explicit goal of Recital 12, to exclude 'simpler systems' and the necessity to include systems that are based on frozen ML models, credit scoring systems based on human experts or symbolic expert systems, imply that two plus signs should not be sufficient, but rather that **three plus signs are required** for a classification as an AI system. It should be noted that this classification does not imply that the relevant system should be considered high-risk or subject to special transparency requirements. Hence, the consequences of qualifying a system as an AI system, as such, are limited.

4. Illustrations

The following applications illustrate characteristics that lead to their classification as either AI systems or non-AI systems. References to commercially available systems are included as examples. Since detailed insights into these systems are not readily available, the descriptions are based on 'best guesses'. Consequently, the classification refers to the presumed characteristics as described, rather than the actual systems mentioned.

Optical Character Recognition (OCR)

OCR technology converts scanned documents and images into machine-readable text by recognising and digitising handwritten or printed characters. The creation of such systems uses large amounts of training data (I ++). The resulting text is perceived by humans as unambiguous at least in the case of printed text (III 0). However, recognising handwritten texts leaves considerable room for interpretation (III+) and may even require interactive online adaptation by the individual writer (II ++).

Chatbot based on a Fixed Open Source LLM (Large Language Model)

A chatbot based on a fixed (open source) large language model uses a model generated from extensive data during the development process (I ++). When generating responses, a pure forward calculation takes place (II 0). The task of conducting a chat conversation typically allows for significant discretion in each generated response (III ++).

Chatbot with API access to an Adaptive LLM (Large Language Model, SaaS)

This type of chatbot uses an API to access a large language model (I ++) that is continuously updated (optimised) based on user feedback (II ++). The task of conducting a chat conversation typically allows for significant discretion in each generated response (III ++).

Search Engine (eg Google, DuckDuckGo, Bing, etc)

Search engines use vast amounts of data obtained from crawling the internet to index and rank websites (PageRank algorithm) (I ++). These indices function (at least internally) as transparent databases that are searched like a telephone directory. Heuristics and (pre-)trained ML methods assist in keywording web pages and interpreting search terms. It is assumed that no online learning or optimisation methods are employed to fulfil the query itself (II 0). The task of delivering the most relevant pages to a search query is not formally defined and involves significant discretion regarding the results (III ++).

Amazon (Recommendation System)

Amazon is a marketplace platform that employs ML to offer personalised product recommendations based on users' previous browsing and purchasing behaviour (I++). These recommendation models are continuously adapted and regularly optimised online through A/B tests based on real user interactions (II++). The task of generating product recommendations specifically customised to the user involves a high degree of discretion (II++).

X (formerly Twitter)

X, formerly known as Twitter, is a social network that uses ML (I ++) to curate each user's personal timeline, identify trends, distinguish fake news and spam, and filter abusive content. The platform incorporates elements of online learning (II ++) to adapt quickly to changing content and user behaviour. These tasks allow for a considerable amount of discretion (III ++).

Music streaming services (eg Spotify)

Spotify is a music streaming service that uses ML to personalise music recommendations based on users' listening habits and preferences (I++). The platform continuously optimises its recommendation models in the background, but only the current (fixed) model is applied during use (II 0). As with other recommender systems, the desired result is highly subjective (III++).

Credit Scores (eg KSV and SCHUFA)

Systems for assessing creditworthiness, such as those of KSV (Austria) and SCHUFA (Germany), analyse extensive amounts of personal financial data to evaluate credit risk and inform lenders about creditworthiness. Traditionally, these systems may use manually programmed rules based on the experience of a few experts (I +) or rely on extensive statistical data (I ++). Although the application of the rules might be straightforward (II 0), the character of the results would allow a human (without usage of automation or AI) significant room for discretion and subjective interpretation (III ++).

Medical Image Diagnostics (eg skin cancer images, CT/MRI images, etc)

Medical image recognition systems utilise machine learning to analyse diagnostic images, such as skin scans or CT and MRI images, assisting doctors in diagnosing and assessing disease states. These systems employ ML models trained on extensive image datasets to perform analysis (I ++). The application of the fixed models to the data is a simple forward calculation (II 0). Like medical examinations, the assessments made by medical image recognition systems would involve a high degree of discretion if human had to decide without support of the AI system (III ++).

Timetable Generators

A timetable Generator uses symbolic scheduling methods to automatically create optimised timetables for educational institutions based on specific requirements and preferences. These systems rely on sophisticated discrete optimisation algorithms (II +) and generally do not require historical data or statistical experience (I 0). In a typical use case, while there is significant flexibility to alter results by inputting design preferences, this discretion space is actively utilised by the user. The automated system adheres strictly to the user's precise specifications to find the desired solution, leaving the automated scheduling solver with no significant discretion (III 0) eventually producing a random draw from a distribution of near equivalent approximate solutions (III +).

Route Planners, Navigation Systems

A simple route planner searches for the shortest distance (II +) on a fixed map with predefined distances (I 0). The system operates based on a database that can be updated as needed (still I 0). The result is uniquely defined (III 0) or an approximate choice from a distribution of close-to equal results (III +).

Adaptive Heating Control Systems

Adaptive heating control systems regulate heating output in buildings to achieve predefined targets and optimise performance according to certain criteria. These systems analyse sensor data such as room temperature, outdoor temperature, window positions and occupancy information to meet the desired targets despite dynamic influences. They use optimisation methods similar to those in reinforcement learning, although typically with very few data points and simple, usually linearised models (II ++). The systems do not require extensive statistical data or historical data (I 0). Due to the small number of factors being optimised, the results can be formally described and verified (III 0).

Adaptive Controllers in Robotics (eg 'Spot')

Adaptive controllers in robotics analyse sensor data to adjust robot operations to the specific requirements of a task or environment. They use dynamic adaptive control systems and reinforcement learning to optimise movements and execution of actions in various dynamic environments (II ++). Large amount of statistical data may have been used is the design (I ++), but also variants with very generic programmes are conceivable (I 0). Although the resulting movement patterns or action plans are formally based on specific goals, predicting whether and how these goals will be achieved is challenging for a human due to the many factors being optimised. Similarly, a human being would exercise a certain degree of discretion when meeting corresponding movement requirements (III +).

Excel Sheets

Excel sheets are spreadsheet-based applications that use a wide range of logic and formula-based functions and integrated optimisation algorithms to analyse and visualise data. These systems allow users to create calculations, charts, and data models based on inputs, defined rules and target formulations. Depending on the type and scope of the data and formulas they contain, Excel applications can perform functions similar to those of other (AI) applications (I 0/+/++). They can depict fully trained ML models and interactively recreate statistical models from data (II 0/+/++). The degree of discretion varies by use case. For instance, a simple invoice form that automates the calculation of discounts and taxes likely involves minimal discretion. However, analysing content and ranking candidates from a list of applicants based on multiple criteria compared to historical candidates would indicate a high degree of discretion (III 0/+/++).

Spelling and Grammar Checkers

Spelling and grammar checkers use machine learning and linguistic rules to recognise errors in texts and suggest corrections. These systems analyse text inputs and compare them with extensive databases of correct word and language patterns to identify potential errors or formal improvements, such as avoiding repeated words (II 0). Spelling and grammar checkers primarily use pattern recognition coded by programmers with expert knowledge, but also employ automatically trained ML models that learn from large language corpora (I ++). As opposed to more advanced tools that make recommendations for formulation improvements (would be III ++), the rules of spelling and grammar (eg Duden) typically leave little or no room for discretion (III 0).

These examples of AI applications could be categorised as follows, based on the descriptions given.

Application	Factor I Development	Factor II Operation	Factor III Outputs	AI System YES/NO
OCR (typewritten text)	++	0	0	NO
OCR (handwriting)	++	0 / ++	+	YES
Chatbot based on a Fixed Open Source LLM	++	0	++	YES
Chatbot with API access to an LLM	++	++	++	YES
Search Engine Google, Duck-Duck-Go, Bing	++	0	++	YES
Recommendation System Amazon	++	++	++	YES
Timeline and filtering at X (formerly Twitter)	++	++	++	YES
Recommendation System Spotify	++	0	++	YES
Credit Scores (eg KSV and SCHUFA)	+ / ++	0	++	YES
Medical Image Diagnostics (eg skin cancer images, CT/MRI images)	++	0	++	YES
Timetable Generator	0	+	+/0	NO
Route Planner, Navigation System	0	+	+	NO
Predictive Maintenance	++	0	++	YES
Adaptive Heating Control Systems	0	++	0	NO
Adaptive Controllers in Robotics (eg 'Spot')	++/0	++	+	YES
Excel Sheet	++/+/0	++/+/0	++/+/0	YES / NO
Spelling and Grammar Checkers	++	0	0	NO

V. Summary

The definition of an 'AI system' in Article 3 No 1 of the AI Act is central to the entire regulatory framework. It is, therefore, surprising that the definition, which is based on the revised OECD definition of November 2023, contains numerous ambiguities, and fails to provide clear guidance on distinguishing AI systems from other IT systems. As a response to the European Commission's public consultation, the ELI proposes a 'Three-Factor Approach' for identifying AI systems: (1) the amount of data or domain-specific empirical knowledge that went into the development and maintenance of the system; (2) the extent to which new know-how is created during the system's operation; and (3) the formal indeterminacy of outputs, ie whether the task at hand is one in which a human would exercise discretion. These three factors should be viewed as interacting within a flexible system, where a strong presence of one factor may offset the weakness or absence of another. Normally, an IT system should be qualified as an AI system where, in the scoring system described above, it receives a total of at least three '+', which also means that at least two of the three factors must be present.

The European Law Institute (ELI) is an independent non-profit organisation established to initiate, conduct and facilitate research, make recommendations and provide practical guidance in the field of European legal development. Building on the wealth of diverse legal traditions, its mission is the quest for better law-making in Europe and the enhancement of European legal integration. By its endeavours, ELI seeks to contribute to the formation of a more vigorous European legal community, integrating the achievements of the various legal cultures, endorsing the value of comparative knowledge, and taking a genuinely pan-European perspective. As such, its work covers all branches of the law: substantive and procedural; private and public.

ELI

EUROPEAN
LAW
INSTITUTE

ISBN: 978-3-9505495-3-9