EPO AI POLICY

Policy on the use of artificial intelligence systems

A Introduction and scope

1. Background

For the last decade, machine learning and artificial intelligence systems (AI) have been developing at a rapid pace, with the potential to fundamentally impact working methods and transform public services.

With the advent of large language models and generative AI in recent years, technology is again becoming a driving force for great change. AI can be used as a tool to improve the effectiveness and efficiency of back-end processes, as well as administrative and legal aspects of the patent grant process ("PGP"), other procedures and beyond (e.g. through its integration in administrative support tasks, financial and human resource administration, procurement, machine translation, image recognition, chatbots, etc).

However, the ability to adopt AI can be hampered by low levels of awareness of the opportunities it offers, or a lack of adequate digital skills, insufficient foundational digital technologies and inadequate digital data. Moreover, AI may pose certain risks, so harnessing its potential in a safe, responsible and sustainable manner will call for tailored measures.

To prepare the European Patent Office (EPO) for wider, secure AI adoption that takes into consideration the safeguards needed to protect individuals' rights and freedoms, this policy provides a clear framework for exploiting opportunities and risk management, which is a prerequisite for implementing and integrating AI at the EPO as part of a robust digital ecosystem that can support future growth.

2. Definition

For the purposes of this policy, "artificial intelligence system" ("AI") means a machine-based system that infers, from the input it receives, how to generate outputs such as predictions, content or recommendations, and to help prepare decisions that may influence physical or virtual environments for explicit or implicit objectives. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment[1]￼

Artificial intelligence encompasses the programming of software systems, often using algorithms, to conduct tasks that have usually required human intelligence in the past.

3. Scope

This policy covers all uses of AI in all business and administrative areas of the EPO.
In addition, this policy should be implemented in agreements with third parties providing services to the EPO, or using our services and tools, by including the appropriate contract terms in such agreements.
In particular, the policy will serve as a basis for those co-operation activities where AI may be of relevance (e.g. AI-supported tools and services). Over the years, the EPO has emerged as a cornerstone of the IP world by sharing numerous services and tools that have helped it and its stakeholders to address technical challenges and harness disruptive technologies. Within the framework of IT co-operation, NPOs are offer support in fully leveraging the technological solutions jointly developed in working groups and projects[2], exploring common practices and future needs and delivering benefits to users across Europe and around the world. Some of these solutions also support international co-operation, tapping into the power of our collective knowledge and expertise to overcome our common challenges.

Lastly, this policy reiterates principles, criteria and safeguards (including monitoring mechanisms) in the areas of IT security and data protection among others, to ensure that rights and freedoms are taken into consideration in the development, deployment and use of AI.

B. The EPO's AI strategy

1. Strategic orientations

- **The EPO is AI-friendly**: The EPO aims to improve the effectiveness, quality, and timeliness of its services and administration. AI is identified as a key enabler for achieving this in a fast, efficient, secure and sustainable way.

- To fully utilise its potential, AI must be integrated in **everyday tools**.

- **The EPO employs a human-centric approach:** The combination of human + AI provides a better result in terms of quality and efficiency than either of them alone, so AI adoption should remain human-centric. With models and solutions evolving rapidly, guidance must be provided in a way that is clear and actionable. While AI allows for greater efficiency and effectiveness in the EPO's operations and administration, final decisions will be taken by humans at the EPO. The accountability and responsibility for such decisions will therefore always remain with the

[1] Definition consistent with Article 2 of the Council of Europe Framework Convention on artificial intelligence and human rights, democracy, and the rule of law of 17 May 2024 and with Article 3(1) of Regulation 2024/1689 laying down harmonised rules on artificial intelligence (AI Act).

[2] For instance, the co-operation initiative "Deep Data and AI" seeks to create a co-operation environment around AI-supported tools. There are plans to develop and deploy AI "functional building blocks". In this context, the national patent offices will have to enforce the principles of this policy.

respective (human) decision-taker, and with the EPO in relationship to third parties, in line with the principles and provisions of the European Patent Convention (EPC). In general, AI will be used in a way that is designed to amplify and augment human abilities.

- **Solutions have been put in place** for pre-search, machine translation, pre-classification and re-classification to improve the quality of products and services. These models were operated and trained with reduced amounts of well-curated data according to data minimisation and accuracy principles and in compliance with data protection rules (DPR).

- **The EPO will continue to leverage AI** to fulfil its mission of creating a more sustainable patent system and a more sustainable society. It will expand on existing AI solutions and continue to explore new areas where AI may be used for the benefit of all stakeholders.

2. Opportunities created by the use of AI

The use of AI opens up unparalleled opportunities:

- **Improved quality:** AI can assist examiners in accessing, analysing and using vast amounts of data – and potentially automate considerable parts of the search – by conducting comprehensive searches for prior art (patent documents or literature, scientific publications, and other relevant sources). In the pre-search phase, we will ensure that the most relevant patent and non-patent literature is identified. This will safeguard high levels of quality and completeness in our search while maintaining timely services.

- Similarly, AI can help to significantly improve the quality of administrative decisions such as financial management, resource allocation or investment decisions. It may also be employed for targeted communication strategies and optimising logistics at the EPO.

- **Improved timeliness:** AI-supported processes can lead to decisions or reporting much more quickly and efficiently than manual processes.

- **Consistency:** AI can ensure consistent practice and administration, e.g. by categorising applications into relevant classes and subclasses, improving the accuracy and consistency of patent classification. The EPO will continue to leverage AI for classification-related tasks, exploring innovative ways of applying this technology to ensure its seamless processing of the ever-mounting volume of prior art. It can also be implemented to support fairer decision-making in HR and procurement or in the evaluation of examination papers (e.g. EQE) by ensuring the consistent application of rules and criteria.

- **Automation and efficiency:** AI has the potential to automate repetitive tasks, leading to increased efficiency and productivity in various areas – including the examination process – by allowing for the use of larger data sets in search and examination, for example. The use of AI to allocate files to the right examiners at the right time – a pillar of our high quality – streamlines the workflow, enhances efficiency and ensures optimal resource allocation, resulting in a more effective and productive examination process.

- **Legal clarity and quality:** AI can assist in the PGP and other legal procedures by providing insights and recommendations based on the entirety of previous decisions, national and international laws, thus facilitating faster and more consistent examination processes. It will enhance the clarity of legal drafting and increase the accuracy of administrative decisions by helping lawyers to verify the consistent and compliant application of rules and criteria.

- **Improved accessibility:** AI can provide real-time translation services, enabling stakeholders to communicate more effectively in different languages and increasing the global reach of the European patent system. The EPO will also harness the power of large language models to process its extensive repository of knowledge, which includes patent manuals, case law, guidelines and many other legal texts. The aim is to seamlessly deliver this valuable information (which will have to be verified for its accuracy) to EPO staff (in particular examiners through the PGP toolset), while ensuring robust data protection safeguards are in place.

- **Informed strategic decision-making:** AI can analyse large volumes of economic and patent data to identify trends, emerging technologies and areas of innovation, as well as providing detailed insights into patent portfolios, which can be useful for policy-making and strategic planning.

3. General principles

The EPO's strategy is to make use of AI where this is considered appropriate and feasible from a legal, compliance and risk perspective.
The EPO will be following principles that have been defined to guide it in making the right decisions and consulting appropriate services:

a. **Respect of fundamental rights:** The EPO will set standards to ensure that no AI system is used that violates fundamental rights, e.g. due process rights, the right to the protection of personal data and privacy and principles of public order. Specifically, the EPO undertakes not to develop, make available, put into service or use an AI system:

- that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective or effect of materially distorting the behaviour of that person or group of persons by appreciably impairing their ability to make an informed decision; thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm

- that exploits the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation, with the objective or effect of materially distorting the behaviour of that person or anyone belonging to that group such that it causes, or is reasonably likely to cause, that person or anyone else significant harm

- for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics ("social scoring"), with the social score leading to either or both of the following:
  - – detrimental or unfavourable treatment of certain natural persons or groups of persons in social contexts that are unrelated to the contexts in which the data was originally generated or collected
  - – detrimental or unfavourable treatment of certain natural persons or groups of persons that is unjustified or disproportionate to their social behaviour or its gravity

- for making risk assessments of natural persons in order to assess or predict the risk of them engaging in a criminal offence or misconduct, based solely on profiling an individual or assessing their personality traits and characteristics. This prohibition does not apply to the use of AI to support human assessment of a person's involvement in a criminal activity or misconduct, which is based on objective and verifiable facts directly linked to that activity or misconduct

- that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage

- to infer a natural person's emotions in the areas of workplace and education institutions, except where the use of AI is intended to be put in place or made available for medical or safety reasons

- that individually categorise natural persons based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation. This excludes the use of AI to confirm that a specific individual is the person they claim to be, to support access management, unblock EPO devices or accounts, recognise a speaker's voice in a video or live conference, manage visitors, or label or filter any lawfully acquired biometric datasets such as images based on biometric data, or to categorise biometric data in administrative investigations and disciplinary procedures.

b. **The EPO will set standards ensuring legal compliance and ethical decision-making, in line with the EPC:** Legal frameworks for managing risks related to the use of AI are still evolving and will be adapted in the future, both on a national, European and international level. The EPO takes note of initiatives such as the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy, and the Rule of Law of 17 May 2024, and the EU AI Act. While none of these instruments are, as such, legally binding on the EPO (due to the Organisation's administrative autonomy under Article 4(1) EPC), the relevant EPO departments will identify the most appropriate use of AI for the EPO, as i.a. set out in this AI policy, and its compliance with the internal legal framework in force at the EPO will be ensured.

c. **The EPO takes a risk-based approach to AI:** The EPO's use of AI will be accompanied by an informed analysis of associated risks, a conscious definition of its appetite for specific risks, as well as the planning and implementation of suitable risk management measures. To this end, Appendix A defines appropriate risk management and impact assessment methodology, which takes into consideration elements such as algorithmic transparency, the risk of bias and dataset quality.

d. **The EPO adopts a full-spectrum approach when using AI:** Leveraging AI will increase efficiency and facilitate more informed decision-making. Consequently, the EPO is looking at how AI can be used in all business areas.

e. **The EPO adopts an agile mindset in dealing with AI:** In the rapidly developing field of AI, new capacities and potential uses are continuously emerging. The EPO will monitor and adapt its approach to and use of AI as appropriate and in response to technical and legal developments.

f. **The EPO follows best practices of data governance**: Specifically in the context of high-risk AI as identified in Appendix A, the EPO will ensure that any AI it develops, deploys and uses is accompanied by high-quality data sets for training, validation and testing. These data sets will be managed properly, considering factors like data collection processes, data preparation, potential biases and data gaps. The data sets used will be relevant, representative, error-free and complete as far as possible. They will also consider the specific context in which AI will be used.

g. **The EPO implements appropriate monitoring and compliance mechanisms**: In order to oversee advancements in the EPO's use of AI, its relevant oversight departments - including Internal Audit, the Data Protection Board and the DPO - will monitor AI's compliance with the EPO's internal rules, including its regulatory framework on information security and the DPR, through data protection audits, for example.

h. **The EPO will set standards to protect individuals from relying on erroneous or inaccurate information provided by AI**: While the EPO seeks to harness the benefits and efficiencies of using AI, AI-generated information is solely intended to assist in decision-making and does not replace human judgment. All responsibility for decisions and actions taken remains with the EPO. Inaccurate or incomplete information generated by AI will not be used as a defence or justification to avoid liability or accountability for decisions made by individuals or bodies acting on behalf of the EPO. For more detailed information on liability, see Section D point 5.

C. Risks of AI

Beyond the opportunities highlighted here, AI also brings certain challenges and risks in terms of data protection, legal compliance, ethical and security, which may impact EPO. The EPO is monitoring these risks closely and will take appropriate action to manage them.

1. Due process and legal risks

Automated decision-making ("ADM") refers to the use of data, machines and algorithms to take decisions, typically with limited human oversight or intervention. AI may assist, support or even supplant decision-making processes. In public administration, this may pose certain legal questions relating to due process principles.

As an example, under Article 24 EPO DPR, data subjects have "*the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning or similarly significantly affecting him or her*". This relates to certain AI-specific risks:

Bias and discrimination: AI algorithms can inadvertently perpetuate or even exacerbate biases present in the data they are trained on. The EPO takes measures to mitigate this risk and ensure that AI systems do not discriminate against any individuals or groups, e.g. through developing criteria to identify high-risk AI (Appendix A), employing diverse datasets, regularly auditing algorithms for bias, and implementing fairness-aware AI techniques.

Responsibility, accountability and oversight: AI systems may make decisions that affect individuals or entities, raising questions about who should be held accountable for those decisions. Establishing mechanisms for accountability and the oversight of AI systems in public administration is crucial to ensure that decisions are made responsibly and in accordance with legal and ethical standards. The Office manages this risk by clearly defining roles and responsibilities for AI governance, implementing a robust governance structure and conducting regular audits. It also provides avenues for redress in cases of harm or injustice.

Transparency: Ensuring the transparency and explainability of AI systems is crucial to meeting due process and data protection requirements. The EPO implements measures to manage this risk, including the adoption of AI systems that provide explanations for their decisions if required.

Copyright risks: AI systems can autonomously generate creative works. Determining the ownership of these works can be complex. AI systems can also be used to create derivative works based on existing copyrighted material. Traditional copyright laws typically grant ownership to human creators, but when AI is involved, questions arise about whether the AI developer, the AI user, or the AI itself should hold the copyright. Furthermore, since AI is trained on vast amounts of data, some of which may be subject to the copyright of third parties, care must be taken that the data used in training, as well as the output of the AI trained with it, does not infringe third party rights. Addressing these copyright challenges calls for the careful consideration of legal frameworks, technological capabilities, and ethical principles to ensure that AI is used in a manner that respects the rights of creators.

2. IT security risks

As a new and powerful technology, AI poses certain security risks, which need to be prudently managed. They include:

Data integrity: AI relies heavily on the quality and integrity of the data it is trained on. Negligence or malfeasance can lead to the manipulation or poisoning of the data used to train AI models, resulting in biased or inaccurate outcomes. Ensuring data integrity is crucial to maintaining the reliability and fairness of AI-driven processes.

Cybersecurity threat vulnerability: AI systems themselves can be vulnerable to cyberattacks. Attackers may exploit vulnerabilities in AI algorithms or manipulate input data to compromise system security.

Use of AI tools in cyber attacks: Additionally, AI-powered systems may be customised and/or used in denial-of-service attacks or other forms of cyber threats, e.g. to manipulate or poison the data used to train AI models, which poses risks to the reliability and fairness of AI-driven processes and the continuity of public administrative operations.

3. Data protection and privacy risks

AI relies on vast amounts of data, including personal data and, in certain cases, special categories of personal data that are sensitive. The processing of personal data by AI has the potential to negatively impact the rights and freedoms of individuals. Examples of specific AI-related risks to be mitigated include:

Identification, aggregation and exposure: AI can pose privacy risks by (i) enabling automated identity linking across various data sources to draw inferences from them, (ii) making use of personal data for purposes other than originally intended, (iii) generating realistic but fake content to spread false or misleading information.

Processing for purposes other than those for which the personal data were collected: AI may use personal data for purposes other than those originally intended by repurposing data, thus impairing respect of the data protection transparency and lawfulness principles.

Distortion: AI can generate realistic but fake content that may be detrimental to the rights and freedoms of individuals through the creation and dissemination of false or misleading information.

4. Ethical considerations

In addition to data protection, legal and security risks, AI also poses certain ethical challenges, which the EPO takes into consideration, including:

Equity and access: The deployment of AI in public administration may exacerbate existing inequalities by disproportionately benefiting certain groups or excluding others. Ensuring equitable access to AI-driven services is essential to prevent widening disparities.

Job displacement and reskilling: The automation of tasks through AI in public administration may lead to job displacement and the need for reskilling programmes to support affected individuals.

D The use of AI

1. Impact assessment and risk management

Prior to the implementation or use of AI, the impact and risks must be appropriately assessed and managed in accordance with this section and the procedure set out in Appendix A.

2. Roles and responsibilities

BIT, together with other EPO units as appropriate, is responsible for analysing all risks related to the implementation or use of AI in its area of competence, as well as for proposing and implementing adequate risk management measures.

The following EPO units must always be involved in any risk assessment:

- The Chief Technology Officer and/or the Chief Information Officer
- the Data Protection Officer
- Information Security

In the case of a high-risk AI system as defined in Appendix A, the following units also may be involved:

- Legal Services
- Patent Law and Procedures
- HR, Employment Law or DG1 depending on the area of competence

3. Compliant and prudent use

AI systems must comply with applicable rules and regulations. The EPO ensures that its use of AI adheres to relevant legal requirements and standards.

EPO users of AI must employ and use AI in compliance with this policy and should be aware of any specific EPO guides that may apply to the use of AI in their area or relate to their specific tasks. In cases where such guides conflict or are incompatible with this policy, the policy prevails.

Users should exercise caution when using AI. With regard to the specific AI they intend to use or are using, users should:

- Be aware of all associated risks
- Be familiar with any instructions or risk management measures in place
- Independently verify any information provided or generated by AI wherever possible
- Never use content generated by AI without modification for any EPO papers or publications
- Attempt to establish and attribute primary sources in AI-compiled content or information
- Avoid inserting personal data in prompts used by AI to carry out its tasks as far as possible

4. Accountability, oversight and continuous improvement

Establishing mechanisms for the accountability and oversight of AI systems in public administration is crucial to ensure that decisions are made responsibly and in accordance with legal and ethical standards. The EPO therefore has a robust governance structure that clearly establishes and communicates decisions on accountability, as well as providing avenues for redress in cases of harm or injustice.

The EPO will continuously update and improve its AI, taking into account emerging best practices, technological developments, and any instances of errors or inaccuracies that may arise in AI's use.

Nothing in this policy affects the principles and rules relating to the responsibility and accountability of EPO employees laid down in the Service Regulations and other internal rules.

5. Liability

The following considerations and principles guide the EPO's approach to liability:

- **Liability for harm to staff or third parties**: In cases where a decision or action is taken, or an omission is made with the assistance of AI to the detriment of a staff member or third party, the EPO will bear responsibility, subject to its internal rules and applicable law. The EPO is committed to taking appropriate steps to rectify any harm caused and ensure that its reliance on AI does not undermine the rights or interests of staff or third parties.

- **AI-assisted services for third parties**: In cases where the EPO provides information or services to third parties using AI, it will take all reasonable measures to ensure the quality and reliability of AI-generated content, but cannot guarantee its absolute accuracy. Third parties who rely on AI-generated information are reminded that such information is provided as a support tool, and not as authoritative or final advice. If information to third parties is provided based on AI-generated content, the EPO will clearly communicate this. However, it accepts no liability for decisions or actions taken by third parties based solely on AI-generated information and encourages independent verification where appropriate.

- **Privileges and immunities:** The Organisation's privileges and immunities, as accorded by Article 8 EPC and the Protocol on Privileges and Immunities, remain unaffected.

E Approval

The EPO's President has approved its Policy on the use of artificial intelligence systems, which is published on the EPO's intranet.

All changes to this document must be approved by the President. New versions of this document will replace all preceding versions.

Appendix A: Impact assessment and risk management

A General impact assessment and risk management provisions

1. Impact and risk assessment process

Prior to their implementation or use, AI systems are subject to an impact and risk assessment process which comprises of the:

a. identification and analysis of the potential uses, benefits and costs of each specific AI system
b. estimation and evaluation of the specific risks that may emerge when AI is used for its intended purpose and under conditions of reasonably foreseeable misuse
c. determination of whether the overall impact of the specific AI system is considered positive, i.e. the benefits identified outweigh negative impact and risks
d. adoption of suitable specific risk management measures

2. Risk management objectives

Risk management measures must aim to:

a. eliminate or minimise risks as far as possible through optimised design and development
b. apply adequate mitigation and control measures to risks that cannot be eliminated
c. provide relevant information and, where appropriate, training to users

3. Residual risk

High-risk AI systems may only be implemented or used if individual residual risks and the overall residual risk level are deemed acceptable after applying appropriate risk management measures and provided that AI is used for its intended purpose or under conditions of reasonably foreseeable misuse.

The criteria for identifying high-risk AI systems detailed in section B below are to be adopted in their assessment, complemented by the AI risk management methodology developed by the relevant EPO departments.

B High-risk AI

1. Criteria for identifying high-risk AI

When determining whether a specific AI system is to be considered as high risk, the following criteria need to be taken into consideration:

a. the intended purpose of the AI system

b. the extent to which an AI system has been used or is likely to be used

c. the nature and amount of the data processed and used by the AI system, in particular whether special categories of personal data are processed

d. the extent to which the AI system acts autonomously and the potential for a human to override a decision or make recommendations that may lead to potential harm

e. the extent to which the use of an AI system has already caused harm, adversely impacted the rights and freedoms of individuals or given rise to major concerns relating to the likelihood of such harm or adverse impact; as demonstrated, for example, by reports or documented allegations submitted to national competent authorities or by other reports, as appropriate

f. the potential extent of such harm or adverse impact, particularly in terms of its intensity and ability to affect multiple persons or disproportionately affect a particular group of persons

g. the extent to which persons who are potentially harmed or suffer an adverse impact depend on the outcome produced by an AI system, particularly if it is not reasonably possible to opt-out of that outcome for practical or legal reasons

h. the extent to which there is an imbalance of power, or the persons who are potentially harmed or suffer an adverse impact are in a vulnerable position in relation to the deployer of an AI system, particularly due to status, authority, knowledge, economic or social circumstances, or age

i. the extent to which the outcome produced with an AI system can be easily corrected or reversed, taking into account the technical solutions available to do so, whereby outcomes that adversely impact health, safety or fundamental rights are not considered to be easily corrigible or reversible

j. the magnitude and likelihood of a benefit for individuals, groups, or society at large from deploying the AI system

2. List of identified areas of high-risk AI

In line with these criteria and in accordance with international standards, the AI systems referred to in the list in this section have been identified as high-risk. The AI systems listed in any of the following areas are considered as high-risk, unless they fall under an exception pursuant to Section B.3 of this Appendix:

a. Biometrics

i. remote biometric identification systems, excluding AI systems intended to be used in biometric verification for the sole purpose of confirming an individual's identity
ii. AI systems intended to be used for biometric categorisation according to sensitive or protected attributes or characteristics based on the inference of those attributes or characteristics
iii. AI systems intended to be used for emotion recognition

b. Education and vocational training

AI systems intended to be used:

i. to determine access or admission or to assign individuals to educational and vocational training institutions at all levels and/or access to professional qualification[3]

[3] E.g. scoring of exams in EQE papers

ii. to evaluate learning outcomes, including when those outcomes are used to steer the learning process of individuals in educational and vocational training institutions at all levels
iii. for the purpose of assessing the appropriate level of education that an individual will receive or will be able to access, in the context of or within educational and vocational training institutions
iv. for monitoring and detecting any prohibited behaviour by students during tests in the context of or within educational and vocational training institutions

c. Employment, worker management and access to self-employment:

AI systems intended to be used:

i. for the recruitment or selection of individuals, in particular to place targeted job advertisements, analyse and filter job applications and to evaluate candidates
ii. to make decisions affecting terms of work-related relationships, the promotion or termination of work-related contractual relationships, to allocate tasks based on individual behaviour or personal traits or characteristics, or to monitor and evaluate the performance and behaviour of persons in such relationships

d. Access to and take-up of essential services and benefits:

AI systems intended to be used to evaluate:

i. the eligibility of individuals for essential benefits and services, including coverage under healthcare insurance provided or paid by the EPO, as well as to grant, reduce, revoke, or reclaim such benefits
ii. eligibility for home loans or other financial assistance, with the exception of AI systems used for the purpose of detecting financial fraud

e. Administrative investigations and disciplinary procedures:

AI systems intended to be used:

i. to assess an individual's risk of becoming the victim of misconduct
ii. as polygraphs or similar tools
ii. to evaluate the reliability of evidence in the course of administrative investigations or disciplinary procedures
iv. for assessing the likelihood that an individual will engage in misconduct not solely based on the profiling of individuals as referred to in Article 3(1)(d) EPO DPR or to assess personality traits and characteristics or past behaviour of individuals or groups
v. AI systems intended to be used for the profiling of individuals as referred to in Article 3(1)(d) EPO DPR in the course of the detection or investigation of misconduct

3. Exceptions: AI systems not considered high-risk

An AI system referred to in Appendix A will not be considered to be high-risk where it does not pose a significant risk of harming the rights and freedoms of individuals, including by not materially influencing the outcome of decision-making.

The AI systems referred to in Appendix A will not be considered to be high-risk when they are intended to:

a. perform a narrow procedural task
b. improve the result of a previously completed human activity
c. detect decision-making patterns or deviations from prior decision-making patterns and are not meant to replace or influence the previously completed human assessment, without proper human review; or
d. perform a preparatory task to an assessment relevant for the purposes of the use cases listed in Appendix A. However, an AI system referred to in Appendix A will always be considered to be high-risk where the AI system performs profiling[4] of data subjects

4. Requirements for high-risk AI systems

a. **Iterative risk management process:** High-risk AI systems are subject to a continuous iterative risk management process run throughout their entire lifecycle that requires regular systematic updating. For high-risk AI systems, the AI impact assessment methodology, prepared by the DPO in collaboration with BIT, will be followed.
b. **Testing:** High-risk AI systems must be tested prior to their implementation or use in a suitable manner for the purposes of identifying the most appropriate risk management measures. Testing ensures that high-risk AI systems perform consistently for their intended purpose.
c. **Training, testing and validation of data:** High-risk AI systems which make use of techniques involving the training of models with data will be developed on the basis of training, validation and testing data sets that meet the highest quality criteria and are subject to appropriate data governance and management practices. Training, validation, and testing data sets must be:
   - relevant
   - representative
   - free of errors
   - complete

They must have the appropriate statistical properties, including as regards the persons or groups of persons on which the high-risk AI system is intended to be used where applicable. The data sets must also take into account, to the extent required by the intended purpose, the characteristics or elements particular to the specific geographical, behavioural or functional setting within which the high-risk AI system is intended to be used. These characteristics may be met at the level of individual data sets or a combination thereof.

d. **Documentation:** The technical documentation of a high-risk AI system must be drawn up before that system is implemented or used and regularly updated.
e. **Transparency:** High-risk AI systems must be designed and developed in such a way as to ensure an appropriate type and degree of transparency, so that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately.

[4] As defined in Article 3(1)(d) EPO DPR, "profiling" means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.

f. **Logs:** High-risk AI systems must be designed and developed with capabilities enabling the automatic recording of events ("logs") while they are operating. Those logging capabilities conform to recognised standards or common specifications and ensure a level of traceability of the AI system's functioning throughout its lifecycle that is appropriate to the intended purpose.

g. **Human oversight:** High-risk AI systems must be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by humans during the period in which the AI system is in use. Human oversight will aim to prevent or minimise the risks to health, safety or fundamental rights that may materialise when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse.

C Limited -risk AI systems

1. **Definition:** Taking into consideration the criteria outlined in Section B.2, AI systems identified as limited risk include any AI technology used:

a. in interactions with humans
b. to generate or manipulate content ("deep fakes")

2. **Transparency:** Limited-risk AI systems must be designed and developed in such a way as to ensure that natural persons are informed that they are interacting with an AI system, unless this is obvious from the circumstances and the context of use.

If the limited-risk AI systems are designed to recognise the user's emotions or employ a biometric categorisation system, any humans exposed to them must be informed accordingly.

Content generated by limited-risk AI systems must be identifiable as such, or explicit information be provided to any humans exposed to them. Text generated by limited-risk AI systems and published with the purpose of informing the public on matters of public interest must be labelled as artificially generated. This also applies to audio and video content constituting deep fakes.

D. Minimal or no risk

Taking into consideration the criteria contained in Section B.2, for any other type of AI system, the general risk management provisions under section A of this Appendix apply.