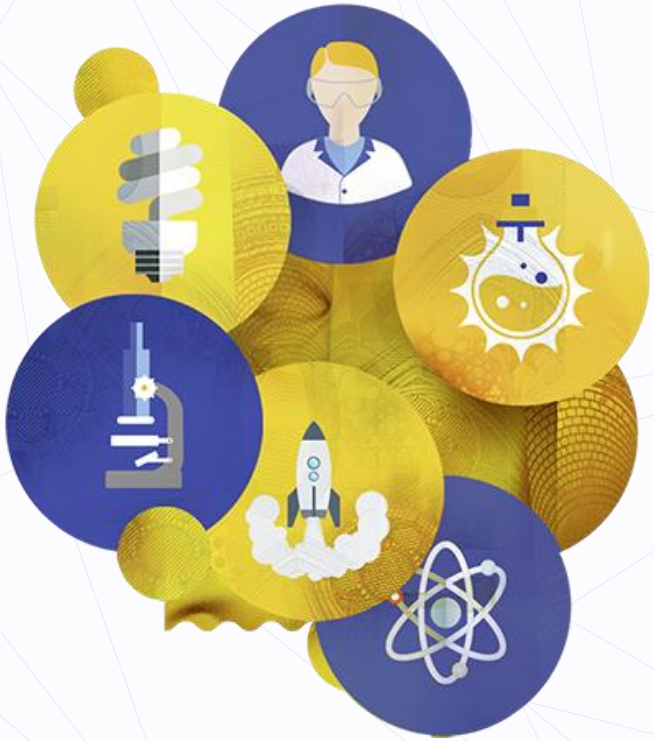
Horizon Europe

MI a tudományban
MI az értékelésben

Németh Edina

Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal



EU AI Strategy – European Approach

In April 2021, the Commission presented its AI package, including:

- its Communication on fostering a European approach to AI;
- a review of the Coordinated Plan on Artificial Intelligence (with EU Member States);
- its Regulatory framework proposal on artificial intelligence and relevant Impact assessment.

https://digital-strategy.ec.europa.eu/en/policies/artificial-intelligence

https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

EU AI Strategy – European Approach

Fostering excellence in AI will strengthen Europe's potential to compete globally.

The EU will achieve this by:

1. enabling the development and uptake of AI in the EU
2. becoming the place where AI thrives from the lab to the market
3. ensuring that AI works for people and is a force for good in society

building strategic leadership in high-impact sectors

EU AI Strategy – European Approach

Resources / funding programmes

Maximising resources and coordinating investments is a critical component of AI excellence. Both the Horizon Europe and Digital Europe programmes will invest €1 billion per year in AI. The Commission will also mobilise additional investments from the private sector and the Member States in order to reach an annual investment volume of €20 billion over the course of the digital decade.

The Recovery and Resilience Facility makes €134 billion available for digital. This will be a game-changer, allowing Europe to amplify its ambitions and become a global leader in developing cutting-edge, trustworthy AI.

Access to high quality data is an essential factor in building high performance, robust AI systems. Initiatives such as the EU Cybersecurity Strategy, the Data act and the Data Governance Act provide the right infrastructure for building such systems.

EU AI Strategy – European Approach

A European approach to trust in AI

Building trustworthy AI will create a safe and innovation-friendly environment for users, developers and deployers.

The Commission has proposed 3 inter-related legal initiatives that will contribute to building trustworthy AI:

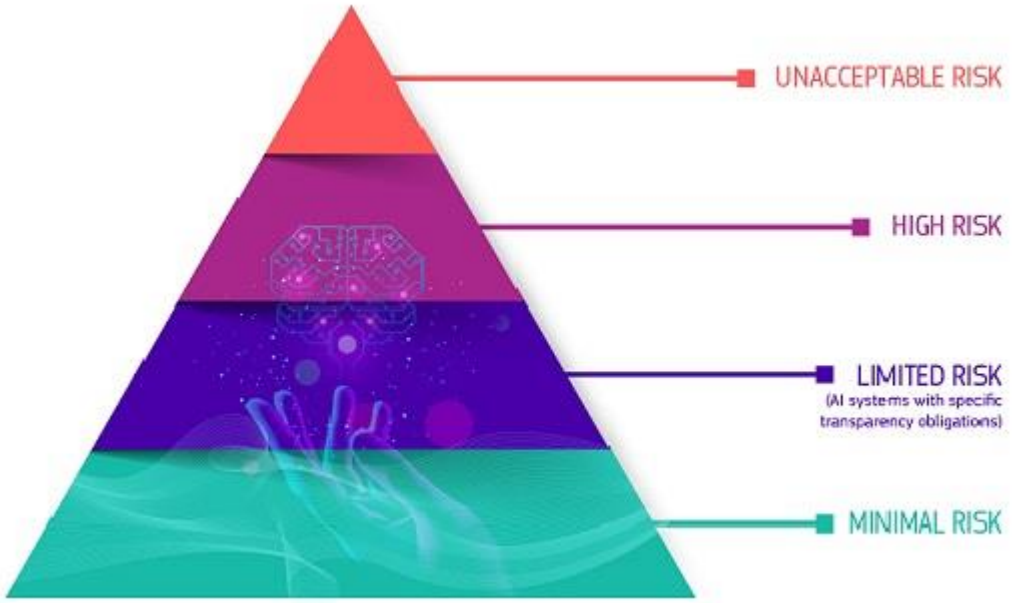
1. European legal framework for AI that upholds fundamental rights and addresses safety risks specific to the AI systems;
2. Civil liability framework - adapting liability rules to the digital age and AI;
3. Revision of sectoral safety legislation (e.g. Machinery Regulation, General Product Safety Directive).

EU AI Strategy – European Approach

European proposal for a legal framework on AI

The Commission aims to address the risks generated by specific uses of AI through a set of complementar proportionate and flexible rules. These rules will alsc provide Europe with a leading role in setting the glol gold standard.

This framework gives AI developers, deployers and users the clarity they need by intervening only in those cases that existing national and EU legislations do not cover. The legal framework for AI, or AI Act, has a clear, easy to understand approach, based on four different levels of risk: minimal risk, high risk, unacceptable risk, and specific transparency risk. It also introduces dedicated rules for general purpose AI models.



European AI Office

The European AI Office will be the centre of AI expertise across the EU. It will play a key role in implementing the AI Act - especially for general-purpose AI - foster the development and use of trustworthy AI, and international cooperation.

sign up to receive updates from the AI Office.

AI in Science

AI in science

Harnessing the power of AI to accelerate discovery and foster innovation : policy brief

The policy brief highlights the following areas for action:

- reducing barriers to adoption of AI in science,
- strengthening the data and compute ecosystem for AI in science,
- identifying strategic R&I investments for the integration of AI in science
- monitoring the impact of AI on research careers,
- addressing AI-related ethical challenges in science.
- communicating, monitoring and evaluation actions to preserve public trust in AI-driven science.

https://op.europa.eu/en/publication-detail/-/publication/094c045c-9e21-11ee-b164-01aa75ed71a1/language-en/format-PDF/source-301545308

Trends in the use of AI in science

A bibliometric analysis

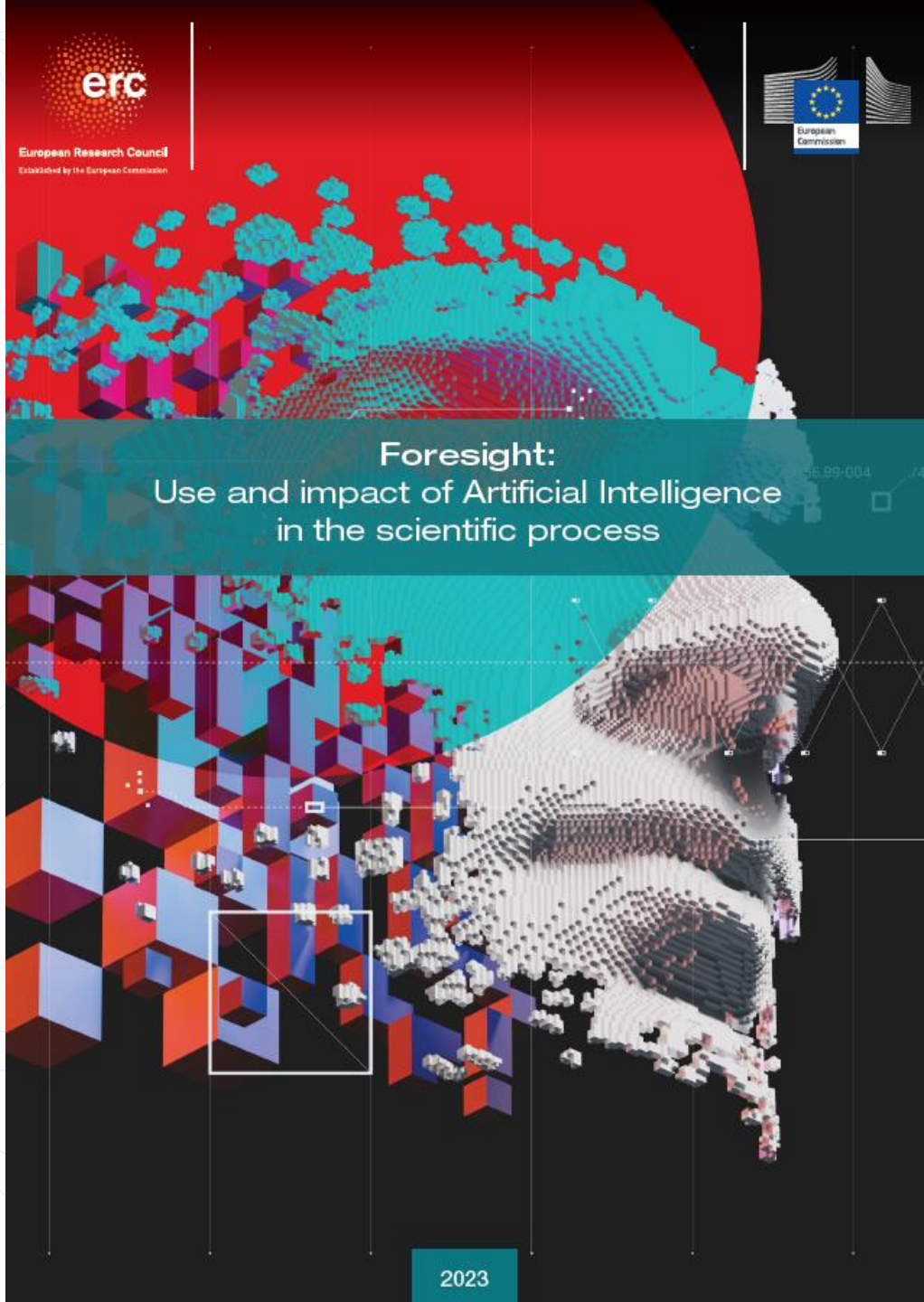
Publication metadata

This paper provides new insights into the role of Artificial Intelligence in scientific research across various domains of science. It is evident that AI as applied to science and research has been growing at a significant pace in recent years, with China leading the way, followed by the EU and the US. Indeed, the EU's performance is found to be highly heterogeneous across its Member States. If current trends persist, the probability of future scientific discoveries being driven primarily by AI applications and tools is set to increase significantly. Failure to keep pace with the development and uptake of AI in science poses important challenges for the EU, including the need to address concerns about strategic autonomy, trends in scientific work and the labour market and, more broadly, productivity and growth.

https://op.europa.eu/en/publication-detail/-/publication/2458267c-08df-11ee-b12e-01aa75ed71a1/language-en/format-PDF/source-287594552

Foresight:

Use and impact of Artificial Intelligence in the scientific process



Not yet delivered

Successful and timely uptake of artificial intelligence in science in the EU

Key downloads

Scoping paper

AI is one of the most disruptive sets of technologies our society has. AI has some characteristics of a 'general purpose technology', being applied in all sectors at high speed and in a disruptive manner, with the promise that it could achieve, and even surpass in some aspects, human abilities.

While science may serve as basis for developing AI-based applications, science is also to be considered as a specific field in which AI can be applied: some of the most complex scientific problems can be tackled with AI technologies (including, but not limited to, data-intensive problems). AI can be a big boost to research,

Scientific advice

Requested by

Margrethe Vestager
Executive Vice-President

Delivery date

April 2024

Supports UN sustainable development goals

Industry, innovation and

https://scientificadvice.eu/events/handover-successful-and-timely-uptake-of-artificial-intelligence-in-science-in-the-eu/

Handover: Successful and timely uptake of Artificial Intelligence in science in the EU

Event details

Start time
15 April 2024 11:45

End time
15 April 2024 12:15

Time zone
CET

Venue
European Commission

Related advice

Successful and timely uptake of artificial

● Online event

On April 15 we will release the advice on how the European Commission can accelerate the responsible take-up of AI in science in the EU.

With the characteristics of a general-purpose technology, artificial intelligence (AI) is one of the most disruptive sets of technologies at the service of research and innovation. It acts as a catalyst for scientific breakthroughs and is rapidly becoming a key instrument in the scientific process in all research areas.

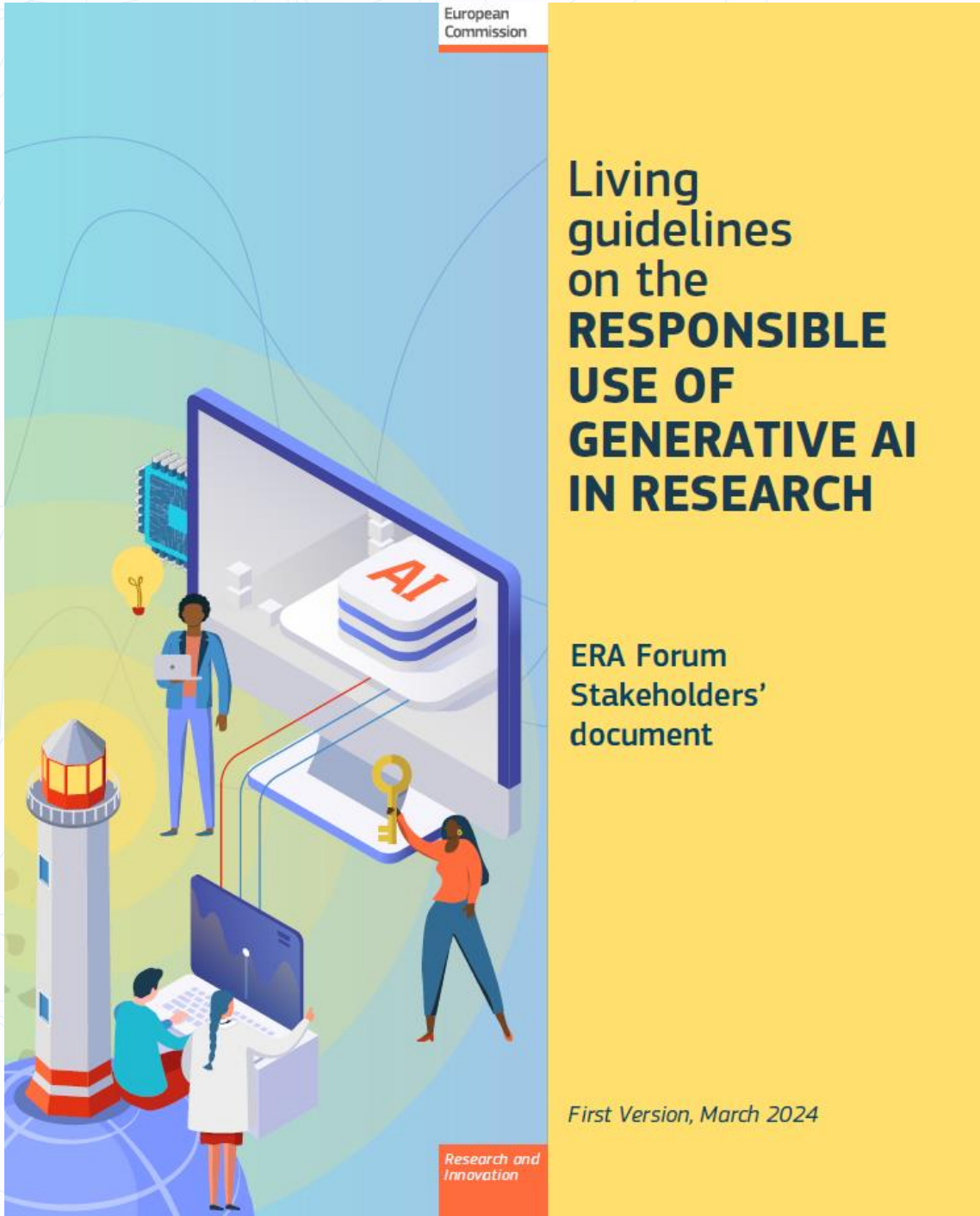
On April 15, we will release the advice on how the European Commission can accelerate the responsible take-up of AI in science in the EU. The release of the evidence review report and policy recommendations will be marked by a ceremony in the presence of EVP Vestager and Commissioner Ivanova.

You can follow the event here.

Feedback on the ERA Forum Stakeholders Document ‘Living Guidelines on the Responsible Use of Generative AI for Research'

AI is one of the most disruptive sets of technologies our society has. AI has some characteristics of a ‘general purpose technology’, being applied in all sectors at high speed and in a disruptive manner, with the promise that it could achieve, and even surpass in some aspects, human abilities.

https://ec.europa.eu/eusurvey/runner/feedback_GenAIResearch



KEY RECOMMENDATIONS

RESEARCHERS should...

- Follow key principles of research integrity, use GenAI transparently and remain ultimately responsible for scientific output.
- Use GenAI preserving privacy, confidentiality, and intellectual property rights on both, inputs and outputs.
- Maintain a critical approach to using GenAI and continuously learn how to use it responsibly to gain and maintain AI literacy.
- Refrain from using GenAI tools in sensitive activities e.g. peer reviews or evaluations.

RESEARCH ORGANISATIONS should...

- Guide the responsible use of GenAI and actively monitor how they develop and use tools.
- Integrate and apply these guidelines, adapting or expanding them when needed.
- Deploy their own GenAI tools to ensure data protection and confidentiality.

FUNDING ORGANISATIONS should...

- Support the responsible use of GenAI in research.
- Use GenAI transparently, ensuring confidentiality and fairness.
- Facilitate the transparent use of GenAI by applicants.

AI Robustness

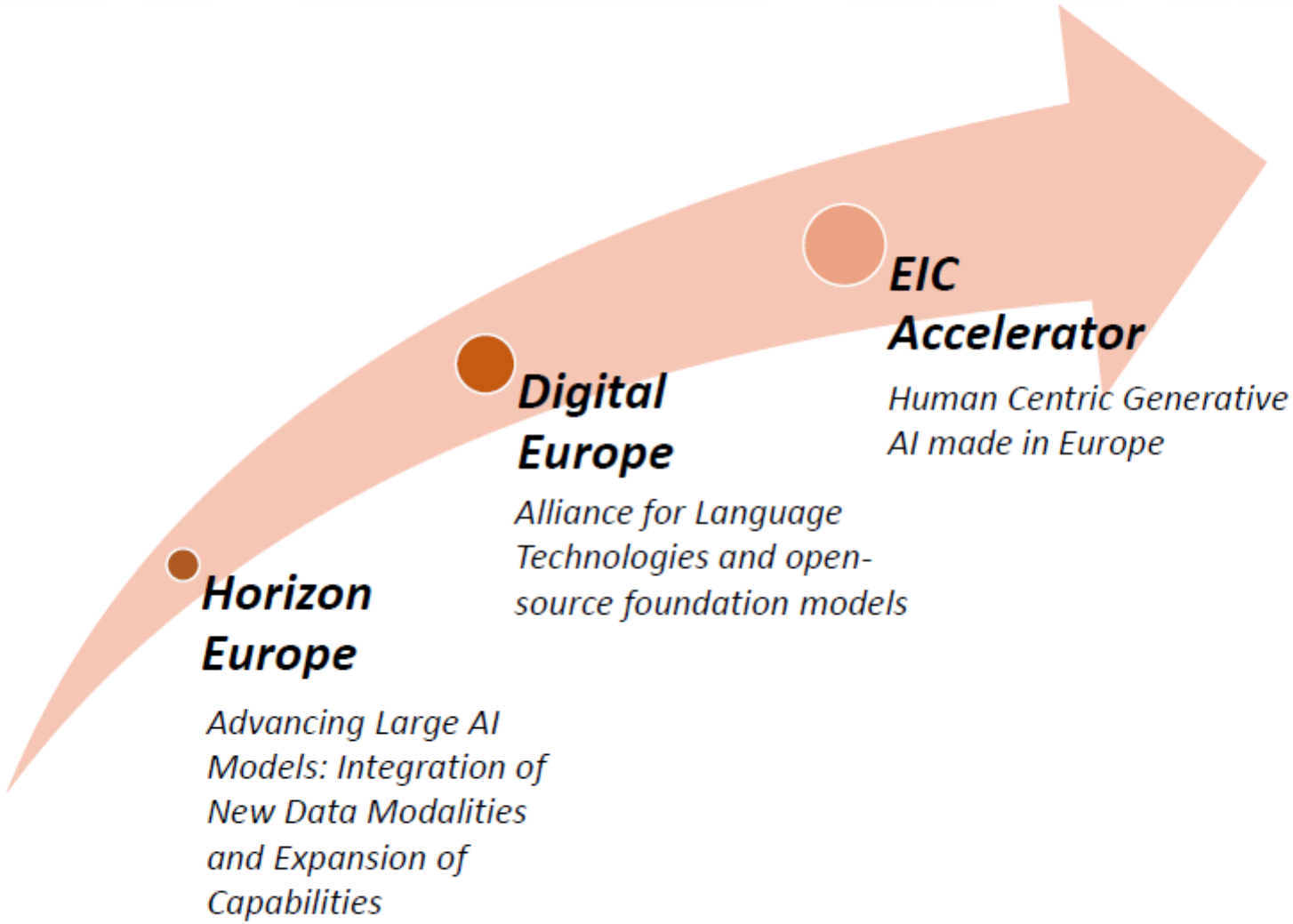
in

Horizon Europe

Proposals & Evaluation

Global AI Framework

Innovation: From the lab to the market



Horizon Europe

Advancing Large AI Models: Integration of New Data Modalities and Expansion of Capabilities (Draft)

Supports the development of foundation AI models:

- Innovative Data Modalities and State-of-the-Art
- Multimodal Models

Explainable and Robust AI *(Draft)*

- Supports the development of AI systems that are more robust, transparent and explainable.

European Innovation Council

Accelerator Challenge: Human Centric Generative AI made in Europe - €50 million

Support the development of foundation language and multimodal 'frontier' models that reach performances at least equivalent to the most powerful state of the art models.

Targeted applicants

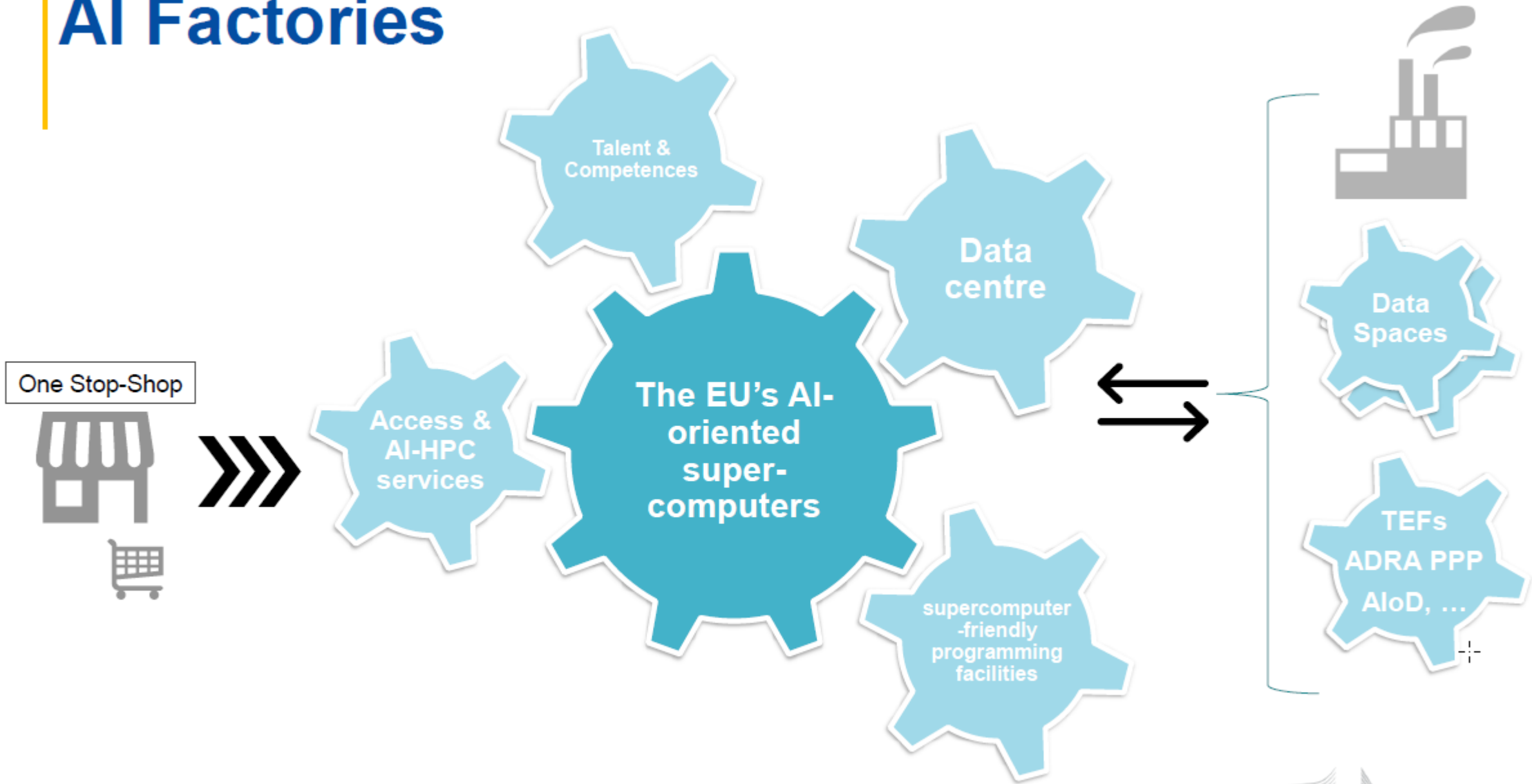
SMEs developing models themselves, and SMEs providing innovative infrastructure, development tools, and critical support.

Communication on boosting startups and innovation in trustworthy artificial intelligence

Two Main Objectives:

- AI Factories
  - Making available HPC computing capacity to facilitate the development of GenAI models
- GenAI4EU
  - Stimulating the development in strategic sectors of novel and innovative applications based on GenAI models and facilitating their uptake.

AI Factories



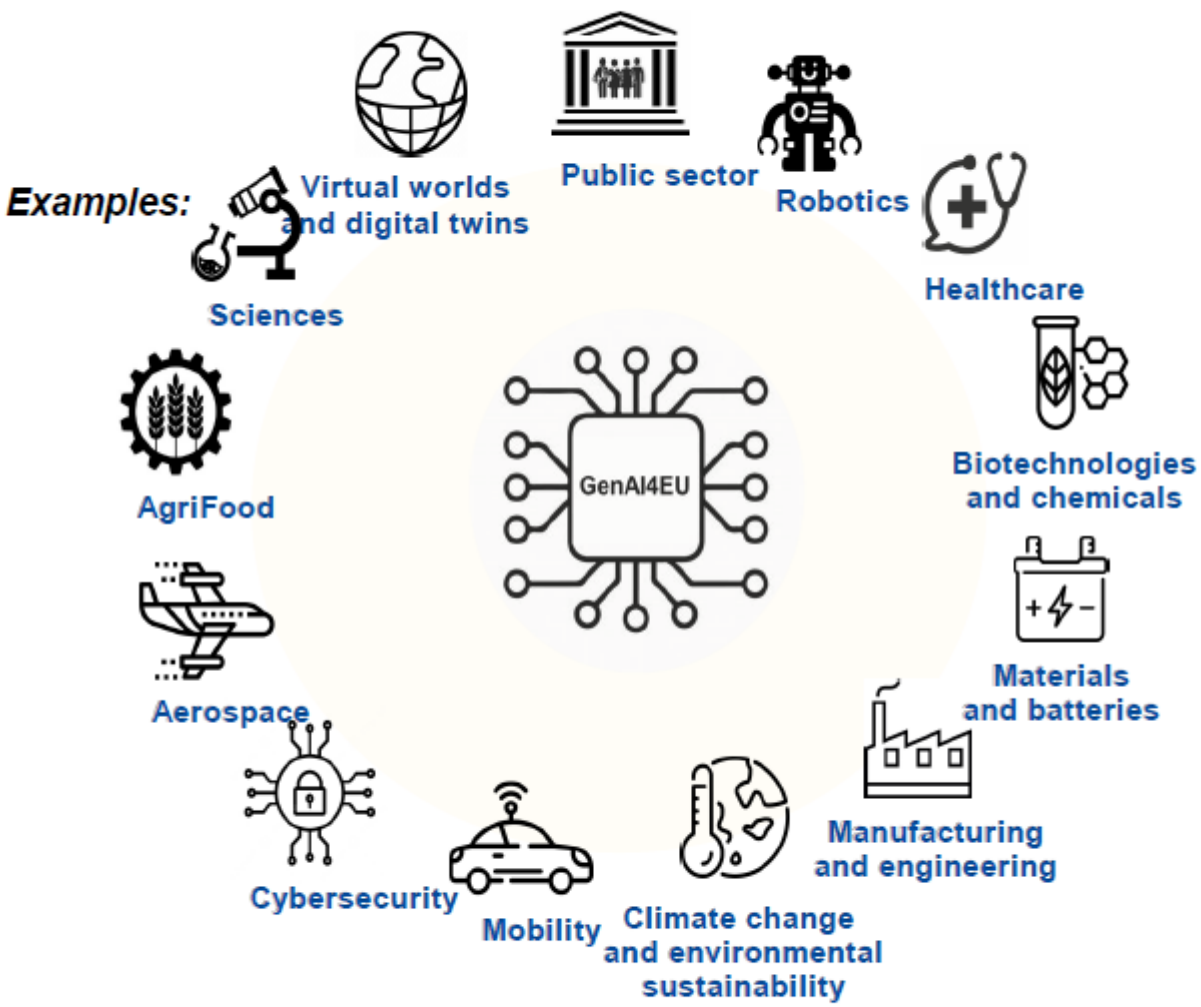
https://digital-strategy.ec.europa.eu/en/policies/ai-factories

European Commission

GenAI4EU Initiative

- **GenAI4EU** initiative to stimulate the widespread uptake of generative AI across the Union's ***fourteen strategic industrial ecosystems.***

- Startups and innovators can work closely with industrial users, attract investments in the Union and have access to the key ingredients of AI - ***data, computing, algorithms and talent.***



AI Robustness

in

Horizon Europe

Proposals & Evaluation

Documentation

- Programme Guide
- Work Programme introduction
- Application form (Part A, Part B)
- Evaluation form
- Evaluator briefing
- Guidelines on ethics by design/operational use for Artificial Intelligence Ethics and Research
- Ethics Guidelines for Trustworthy Artificial Intelligence (AI)
- factsheet on gender and intersectional bias in AI
- …. and of course the call topic

- Tools … CAP AI, Ethics by design - roles & responsibilities

IDEAL-IST

Programme Guide

Gendered Innovation

A full policy report has been prepared and is available to support applicants. Entitled *Gendered Innovations 2: How inclusive analysis contributes to research and innovation* and publicly released by the European Commission on 25 November 2020, it is available here, through the Europa website dedicated to gender equality policy in R&I.

The report contains: full definitions of terms; both general and field-specific methods for sex analysis, gender analysis and intersectional approaches; fifteen case studies covering health, climate change, energy, agriculture, urban planning, waste management, transport, artificial intelligence (AI) and digital technologies, taxation, venture funding, as well as COVID-19; and policy recommendations to address the global challenges, targeted impacts and key R&I orientations of the six Horizon Europe Clusters, as well as Mission Areas, and European partnerships.

More information and examples on how to integrate the gender dimension into R&I content in different fields of research and innovation may be found here:

- Website developed by the EU-supported Expert Group on Gendered Innovations, featuring latest material presented in the 2020 EC policy report *Gendered Innovations 2: How inclusive analysis contributes to research and innovation*, as well as previous case studies developed through EC support
- Factsheets:
  - factsheet summarising the EC policy report's contents
  - factsheet on the impact of sex and gender in the COVID-19 pandemic
  - factsheet on gender and intersectional bias in AI
  - factsheet on general provisions for gender equality under Horizon Europe

Programme Guide

Ethics Review

The ethics review covers issues as:

- human rights and protection of human beings
- animal protection and welfare
- data protection and privacy
- health and safety
- environmental protection
- artificial intelligence

Programme Guide

Key Digital Technologies

Due diligence is required regarding the trustworthiness of all artificial intelligence-based systems or techniques used or developed in projects funded under the Horizon Europe Programme. Wherever appropriate, AI-based systems or techniques must be developed in a safe, secure and responsible manner, with a clear identification of and preventative approach to risks.

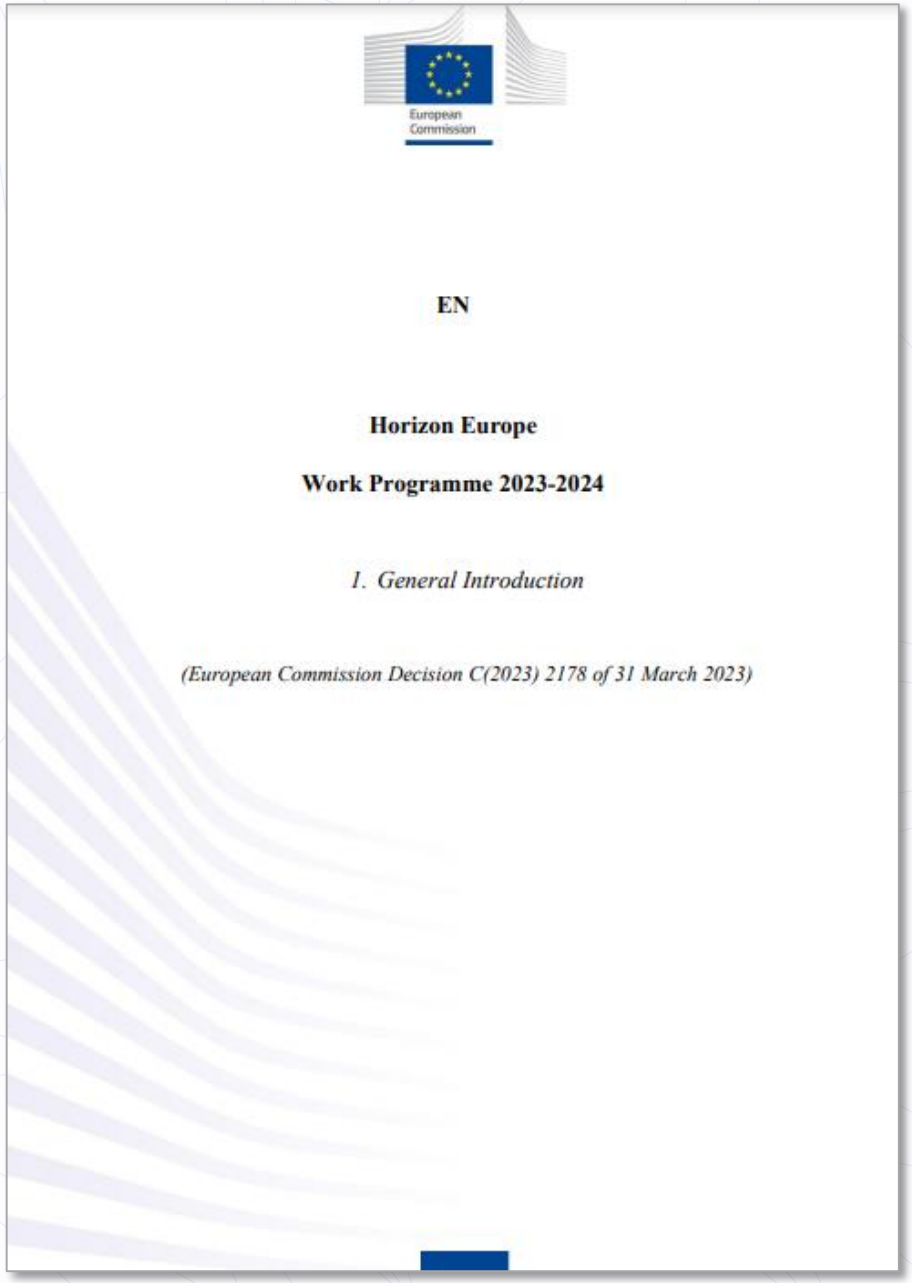
To a degree matching the type of research being proposed (from basic to precompetitive) and as appropriate, AI-based systems or techniques should be, or be developed to become (implicitly or explicitly contributing to one or several of the following objectives):

- **technically robust**, accurate and reproducible, and able to deal with and inform about possible failures, inaccuracies and errors, proportionate to the assessed risk posed by the AI-based system or technique
- **socially robust**, in that they duly consider the context and environment in which they operate
- **reliable** and to function as intended, minimising unintentional and unexpected harm, preventing unacceptable harm and safeguarding the physical and mental integrity of humans
- able to provide a suitable **explanation of its decision-making** process, whenever an AI-based system can have a significant impact on people's lives.

Work Programme Intro

Four impact areas

To promote industrial leadership in **key and emerging technologies that work for people**, the co-programmed Partnership on Artificial Intelligence, Data and Robotics will drive the development of **human-centric, trustworthy, safe and robust technologies** that will boost new markets and applications and that are compatible with **Europe's ethical standards** and values. A dedicated action will also examine **humanistic deployment of artificial intelligence and related technologies**.


European Commission

EN

Horizon Europe

Work Programme 2023-2024

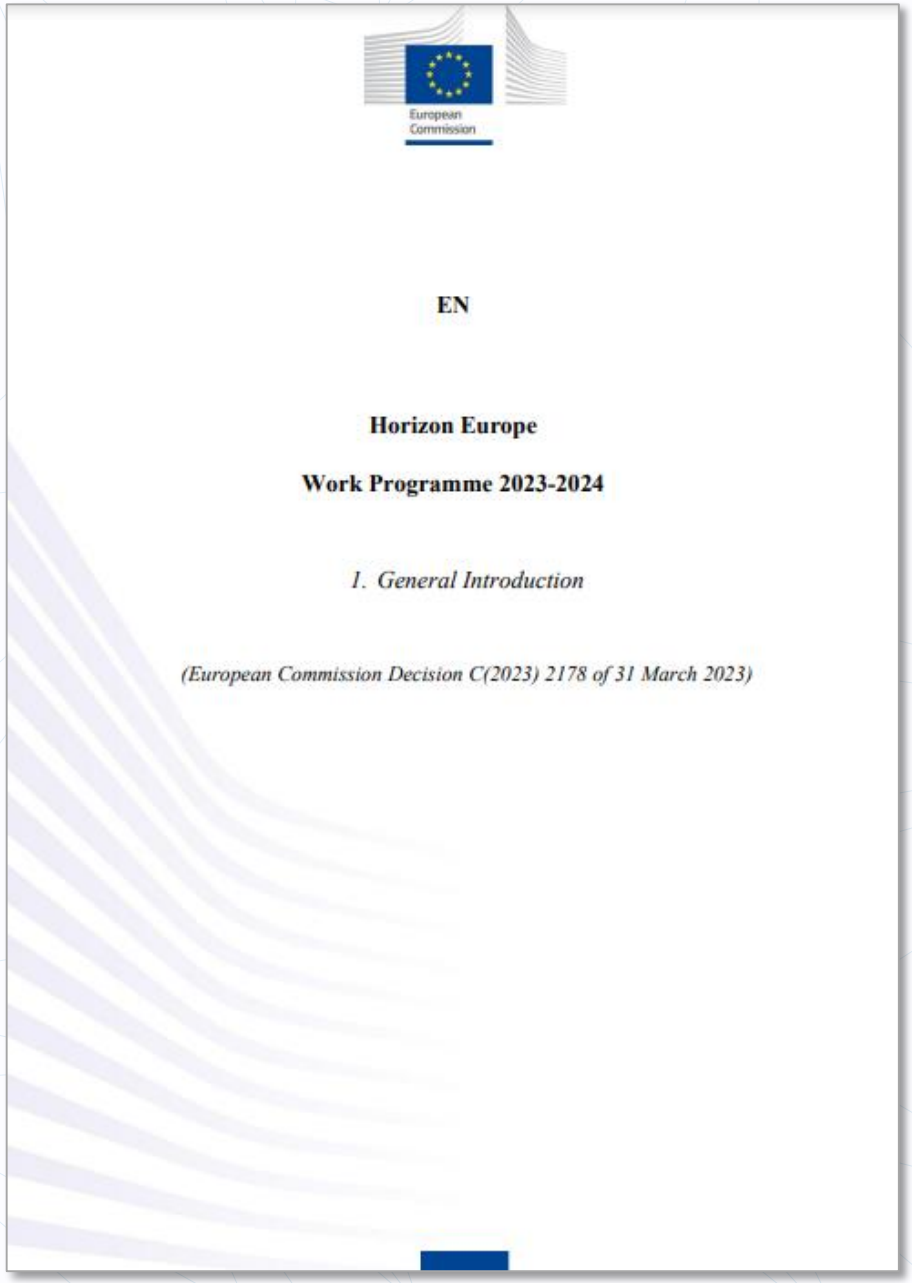
1. General Introduction

(European Commission Decision C(2023) 2178 of 31 March 2023)

Work Programme Intro

Trustworthy technologies

All projects supported by this work programme will be in line with **EU values** and adhere to the highest **ethics and integrity standards**. Horizon Europe is spearheading the artificial intelligence **ethics by design** agenda. Due diligence will be required to make sure all AI-based systems or techniques used or developed will be **trustworthy: ethical, lawful and robust, with particular attention to safety, accuracy, reliability and explainability**.



Application forms

< Budget | Ethics & Security | Other questions >

Table of contents | Validate form | Save form | Save & exit form

Administrative forms

Proposal ID **SEP-211003912**

Acronym **123424412**

Does this activity involve low and/or lower middle income countries, (if yes, detail the benefit-sharing actions planned in the self-assessment)	○ Yes ◉ No	
Could the situation in the country put the individuals taking part in the activity at risk?	○ Yes ◉ No	
7. Environment, Health and Safety		**Page**
Does this activity involve the use of substances or processes that may cause harm to the environment, to animals or plants.(during the implementation of the activity or further to the use of the results, as a possible impact) ?	○ Yes ◉ No	
Does this activity deal with endangered fauna and/or flora / protected areas?	○ Yes ◉ No	
Does this activity involve the use of substances or processes that may cause harm to humans, including those performing the activity.(during the implementation of the activity or further to the use of the results, as a possible impact) ?	○ Yes ◉ No	
8. Artificial Intelligence		**Page**
Does this activity involve the development, deployment and/or use of Artificial Intelligence-based systems?	◉ Yes ○ No	
9. Other Ethics Issues		**Page**
Are there any other ethics issues that should be taken into consideration?	○ Yes ◉ No	

I confirm that I have taken into account all ethics issues above and that, if any ethics issues apply, I will complete the ethics self-assessment as described in the guidelines How to Complete your Ethics Self-Assessment ☐ ?

Application forms

1. Excellence #@REL-EVA-RE@#

#§PRJ-OBJ-PO§#

1.2 Methodology #@CON-MET-CM@# #@COM-PLE-CP@# *[e.g. 14 pages]*

- Describe and explain the overall methodology, including the concepts, models and assumptions that underpin your work. Explain how this will enable you to deliver your project's objectives. Refer to any important challenges you may have identified in the chosen methodology and how you intend to overcome them. *[e.g. 10 pages]*

⚠ *This section should be presented as a narrative. The detailed tasks and work packages are described below under 'Implementation'.*

⚠ *Where relevant, include how the project methodology complies with the 'do no significant harm' principle as per Article 17 of Regulation (EU) No 2020/852 on the establishment of a framework to facilitate sustainable investment (i.e. the so-called 'EU Taxonomy Regulation'). This means that the methodology is designed in a way it is not significantly harming any of the six environmental objectives of the EU Taxonomy Regulation.*

⚠ *If you plan to use, develop and/or deploy artificial intellingence (AI) based systems and/or techniques you must demonstrate their technical robustness. AI-based systems or techniques should be, or be developed to become:*

- *technically robust, accurate and reproducible, and able to deal with and inform about possible failures, inaccuracies and errors, proportionate to the assessed risk they pose*
- *socially robust, in that they duly consider the context and environment in which they operate*
- *reliable and function as intended, minimizing unintentional and unexpected harm, preventing unacceptable harm and safeguarding the physical and mental integrity of humans*
- *able to provide a suitable explanation of their decision-making processes, whenever they can have a significant impact on people's lives.*

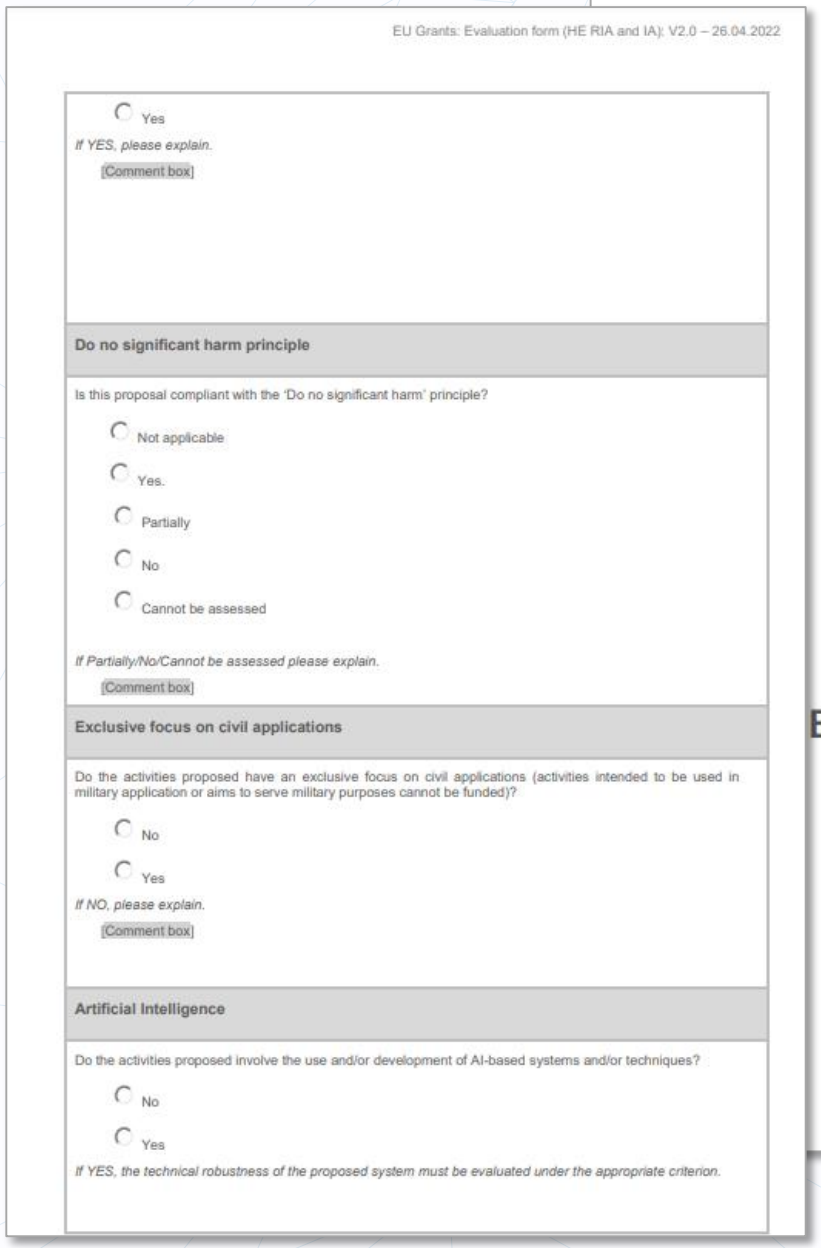
Evaluation forms

Artificial Intelligence

Do the activities proposed involve the use and/or development of AI-based systems and/or techniques?

- No
- Yes

If YES, the technical robustness of the proposed system must be evaluated under the appropriate criterion.



Yes

If YES, please explain.

[Comment box]

Do no significant harm principle

Is this proposal compliant with the 'Do no significant harm' principle?

Not applicable

Yes.

Partially

No

Cannot be assessed

If Partially/No/Cannot be assessed please explain.

[Comment box]

Exclusive focus on civil applications

Do the activities proposed have an exclusive focus on civil applications (activities intended to be used in military application or aims to serve military purposes cannot be funded)?

No

Yes

If NO, please explain.

[Comment box]

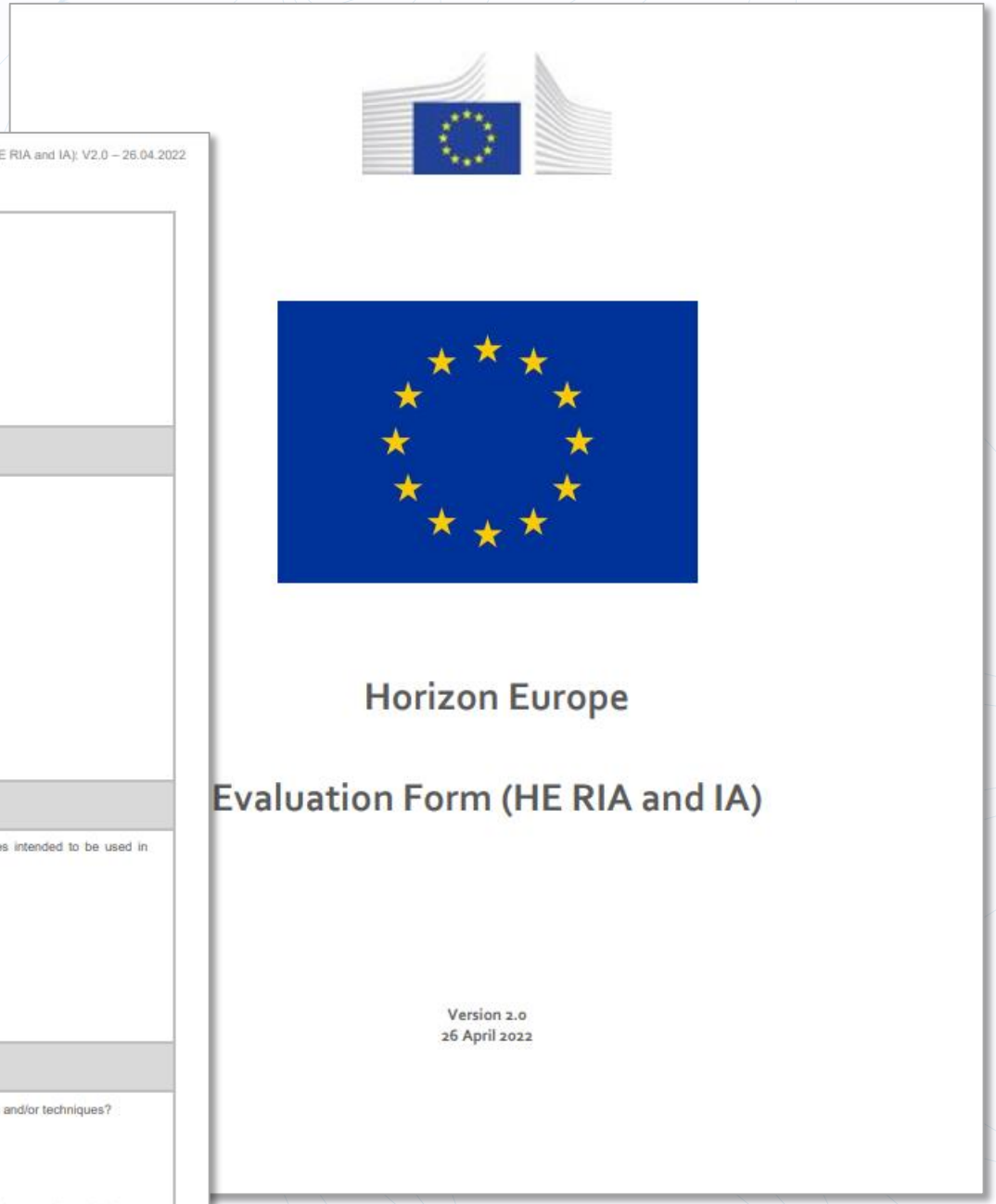
Artificial Intelligence

Do the activities proposed involve the use and/or development of AI-based systems and/or techniques?

No

Yes

If YES, the technical robustness of the proposed system must be evaluated under the appropriate criterion.



Evaluator briefing

Additional questions in the evaluation form

Evaluation form includes:

- Main part with the three **evaluation criteria** where you give comments and scores
- **Additional questions:** The evaluators are asked to take a position on additional questions linked to the selection procedure or policy considerations.

Additional questions in Horizon Europe evaluations

- Scope of the application
- Additional funding
- Use of human embryonic stem cells (hESC)
- Use of human embryos (hE)
- Activities not eligible for funding
- Exclusive focus on civil applications
- Do not significant harm principle
- Artificial Intelligence

European Commission

Artificial intelligence

- Experts must answer an additional question as part of their individual evaluations on whether the activities proposed involve the **use and/or development of AI-based systems and/or techniques**.
- If you answer 'yes' to this question, you must **assess the technical robustness** of the proposed AI-system as part of the excellence criterion (if applicable).
- In addition, your answer to this question will help us to with the **proper follow-up** of any aspects related to **Artificial Intelligence** in projects funded under Horizon Europe.

(*) Technical robustness refers to technical aspects of AI systems and development, including resilience to attack and security, fullback plan and general safety, accuracy, reliability and reproducibility.

AI-based systems or techniques should be, or be developed to become:

- **Technically robust**, **accurate and reproducible**, and able to deal with and inform about possible failures, inaccuracies and errors, proportionate to the assessed risk posed by the AI-based system or technique.
- **Socially robust**, in that they duly consider the context and environment in which they operate.
- **Reliable and function as intended**, minimizing unintentional and unexpected harm, preventing unacceptable harm and safeguarding the physical and mental integrity of humans.
- Able to provide a suitable explanation of its **decision-making process**, whenever an AI-based system can have a significant impact on people's lives.

ion

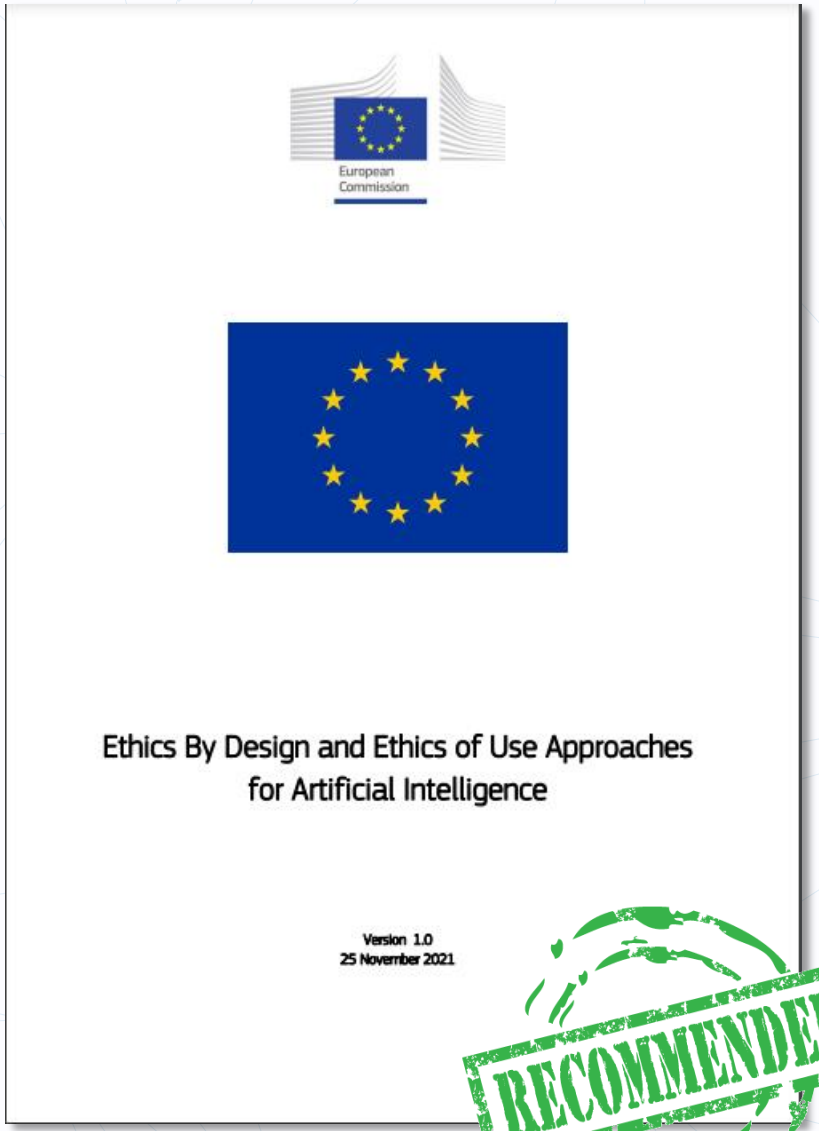
Guidance - Ethics by Design

Guidance for adopting an ethically-focused approach while designing, developing, and deploying and/or using AI based solutions.

It explains the ethical principles which AI systems must support and discusses the key characteristics that an AI-based system/ applications must have in order to preserve and promote:

- respect for human agency;
- privacy, personal data protection and data governance;
- fairness;
- individual, social, and environmental well-being;
- transparency;
- accountability and oversight.

Furthermore, it details specific tasks which must be undertaken in order to produce an AI which possess these characteristics.



Ethical Principles and Requirements

There are six general ethical principles that any AI system must preserve and protect based on fundamental rights as enshrined in the Charter of Fundamental Rights of the European Union (EU Charter), and in relevant international human rights law:

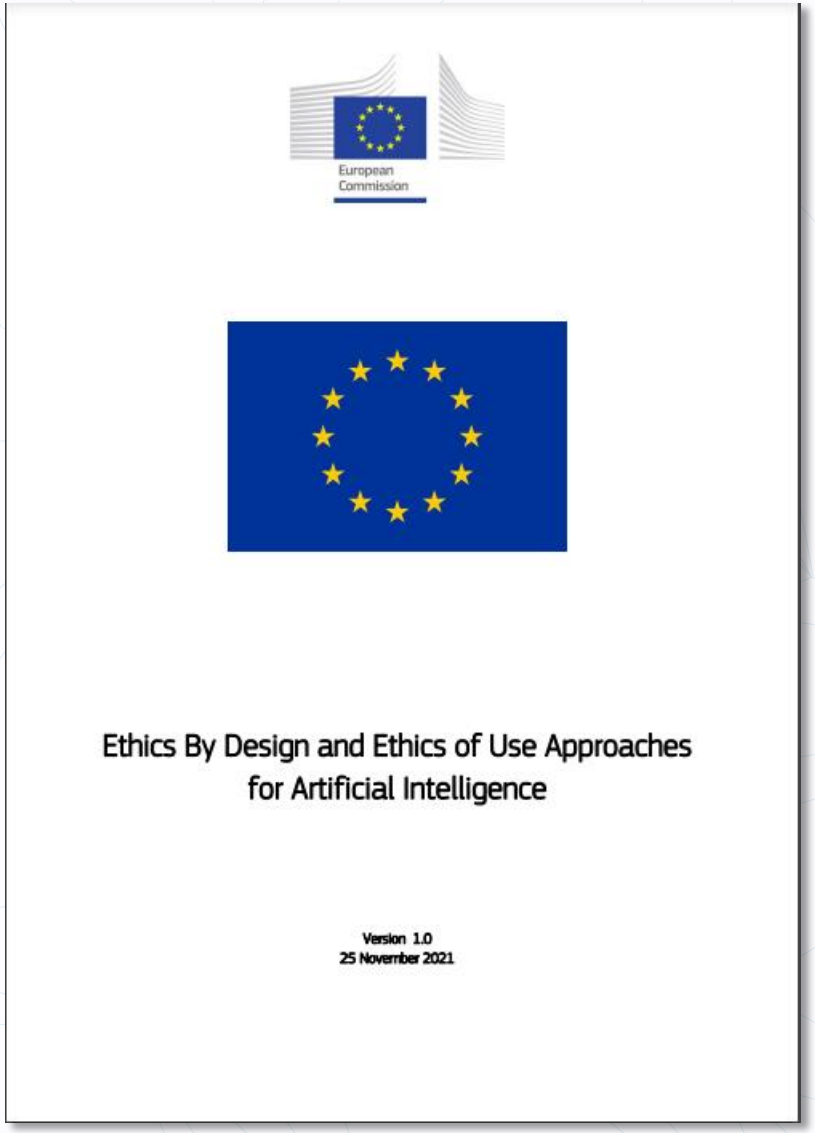
1. **Respect for Human Agency:** human beings must be respected to make their own decisions and carry out their own actions. Respect for human agency encapsulates three more specific principles, which define fundamental human rights: autonomy, dignity and freedom.
2. **Privacy and Data governance:** people have the right to privacy and data protection and these should be respected at all times;
3. **Fairness:** people should be given equal rights and opportunities and should not be advantaged or disadvantaged undeservedly;
4. **Individual, Social and Environmental Well-being:** AI systems should contribute to, and not harm, individual, social and environmental wellbeing;
5. **Transparency:** the purpose, inputs and operations of AI programs should be knowable and understandable to its stakeholders;
6. **Accountability and Oversight:** humans should be able to understand, supervise and control the design and operation of AI based systems, and the actors involved in their development or operation should take respo

Ethics by Design

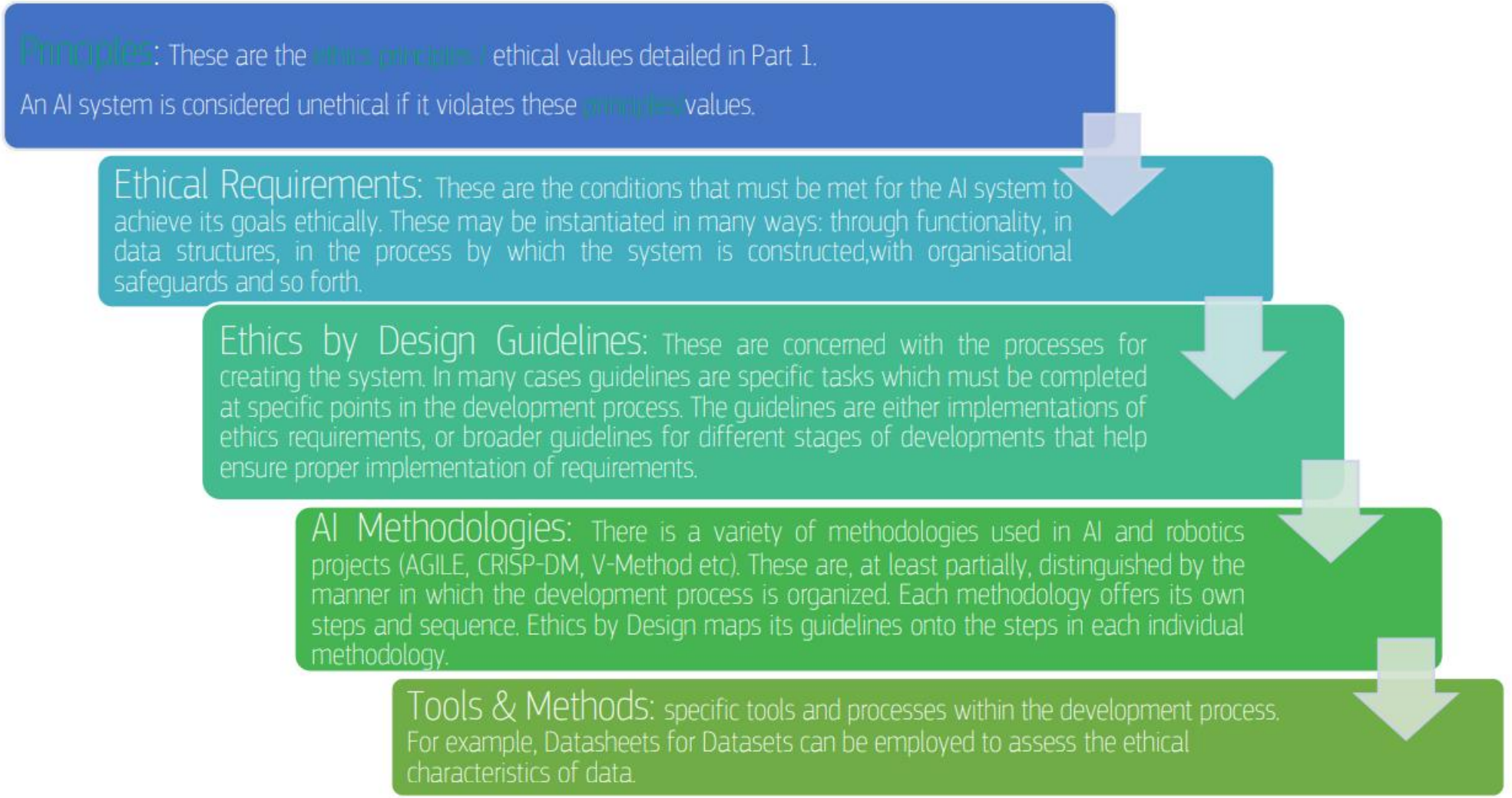
- **Part 1: Principles and requirements:** This part defines the ethical principles that AI systems should adhere to and derives requirements for their development;

- **Part 2: Practical steps for applying Ethics by Design in AI development:** This section explains the Ethics by Design concept and relates it to a generic model for the development of AI systems. It defines the actions to be taken at different stages in the AI development in order to adhere to the ethics principles and requirements listed in Part 1;

- **Part 3: Ethical deployment and use** presents guidelines for deploying or using AI in an ethically responsible manner.



5 layer model of ethical design



Ethical Requirements for AI & Robotics Systems

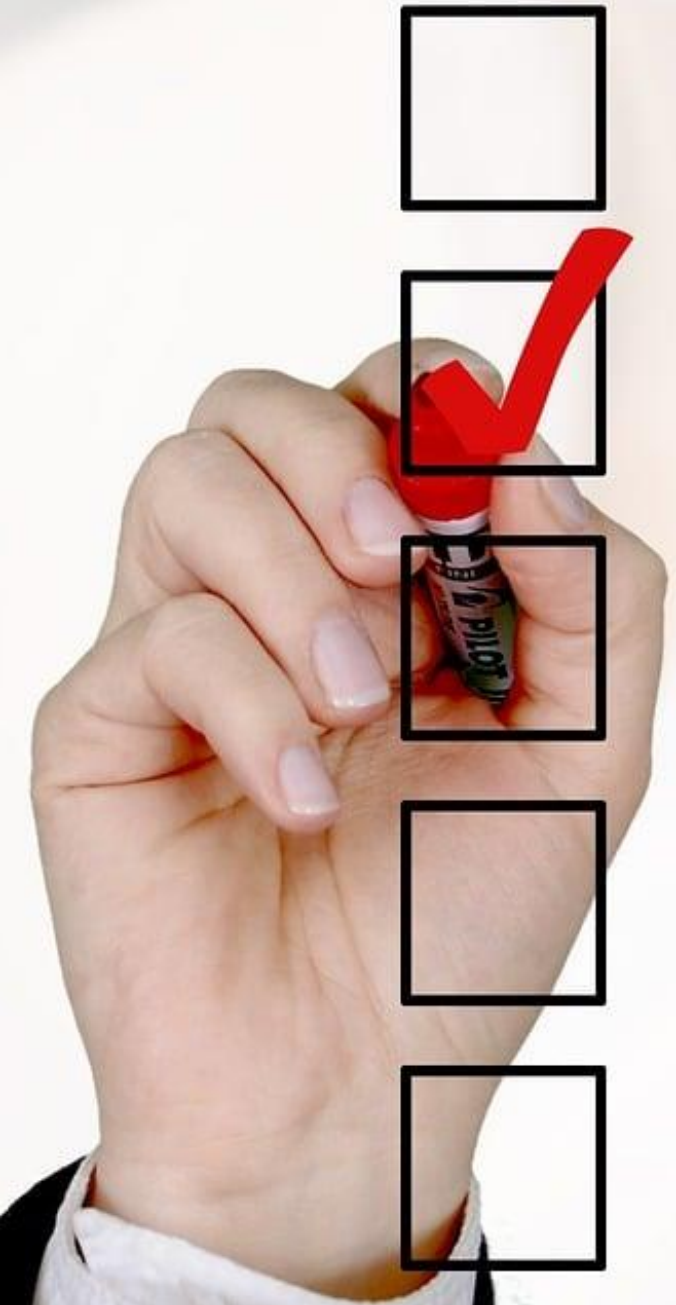
The main ethical requirements for AI and robotics systems above can be summarised as:

- AI systems must not negatively affect human autonomy, freedom or dignity.
- AI systems must not violate the right to privacy and to personal data protection. They MUST use data which is necessary, non-biased, representative and accurate.
- AI systems must be developed with an inclusive fair, and non-discriminatory agenda.
- Steps must be taken to ensure that AI systems do not cause individual, social or environmental harm, rely on harmful technologies, influence others to act in ways which cause harm or lend themselves to function creeps.
- AI systems should be as transparent as possible to their stakeholders and to their end-users.
- Human oversight and accountability are required to ensure conformance to these principles and address non-compliance.

Ethics by Design – Generic Development Model

The six tasks in the generic model are:

1. **Specification of objectives:** The determination of what the system is for and what it should be capable of doing.
2. **Specification of requirements:** Development of technical and non-technical requirements for building the system, including initial determination of required resources, together with an initial risk assessment and cost-benefit analysis, resulting in a design plan.
3. **High-level design:** Development of a high-level architecture. This is sometimes preceded by the development of a conceptual model.
4. **Data collection and preparation:** Collection, verification, cleaning and integration of data.
5. **Detailed design and development:** The actual construction of a fully working system.
6. **Testing and evaluation:** Testing and evaluation of the system.



Annex I Checklist: Specification of Objectives against Ethical Requirements

This checklist is a supporting tool and does not constitute an exhaustive list of all ethics requirement that ma
be applicable to the development of each specific AI system. It has to be used in conjunction with Part 1-3
the current guidelines and applied to a degree matching the type of AI system and the research being propos
(from basic to precompetitive).

Specification of Objectives against Ethical Requirements	**Yes**	**No (how potent risks will mitigated?)**
Respect for Human Agency		
End-users and others affected by the AI system are not deprived of abilities to make all decisions about their own lives, have basic freedoms taken away from them,		
End-users and others affected by the AI system are not subordinated, coerced, deceived, manipulated, objectified or dehumanized, nor is attachment or addiction to the system and its operations being stimulated.		
The system does not autonomously make decisions about vital issues that are normally decided by humans by means of free personal choices or collective deliberations or similarly significantly affects individuals,		
The system is designed in a way that give system operators and, as much as possible, end-users the ability to control, direct and intervene in basic operations of the system (when relevant)		
Privacy & Data Governance		
The system processes data in line with the requirements for lawfulness, fairness and transparency set in the national and EU data protection legal framework and the reasonable expectations of the data subjects.		
Technical and organisational measures are in place to safeguard the rights of data subjects (through measures such as anonymization, pseudonymisation, encryption, and aggregation).		
There are security measures in place to prevent data breaches and leakages (such as mechanisms for logging data access and data modification).		
Fairness		
The system is designed to avoid algorithmic bias, in input data, modelling and algorithm design. The system is designed to avoid historical and selection bias in data collection, representation and measurement bias in algorithmic training, aggregation and evaluation bias in modelling and automation bias in deployment		
The system is designed so that it can be used different types of end-users with different abilities (whenever possible/relevant)		
The system does not have negative social impacts on relevant groups, including impacts other than those resulting from algorithmic bias or lack of universal accessibility,		
Individual, and Social and Environmental Well-being		
The AI system takes the welfare of all stakeholders into account and do not unduly or unfairly reduce/undermine their well-being		
The AI system is mindful of principles of environmental sustainability, both regarding the system itself and the supply chain to which it connects (when relevant)		
The AI system does not have the potential to negatively impact the quality of communication, social interaction, information, democratic processes, and social relations (when relevant)		
The system does not reduce safety and integrity in the workplace and complies with the relevant health and safety and employment regulations		
Transparency		
The end-users are aware that they are interacting with an AI system		
The purpose, capabilities, limitations, benefits and risks of the AI system and of the decisions conveyed are openly communicated to and understood by end-users and other stakeholders along with its possible consequences		
People can audit, query, dispute, seek to change or object to AI or robotics activities (when applicable)		
The AI system enables traceability during its entire lifecycle, from initial design to post-deployment evaluation and audit		
The system offers details about how decisions are taken and on which reasons these were based (when relevant and possible)		
The system keeps records of the decisions made (when relevant)		
Accountability & Oversight		
The system provides details of how potential ethically and socially undesirable effects will be detected, stopped, and prevented from reoccurring.		
The AI system allows for human oversight during the entire life-cycle of the project /regarding their decision cycles and operation (when relevant)		

Ethics Guidelines for Trustworthy AI

Ethics Principles - Key guidance:

- Develop, deploy and use AI systems in a way that adheres to the ethical principles of: respect for human autonomy, prevention of harm, fairness and explicability. Acknowledge and address the potential tensions between these principles.
- Pay particular attention to situations involving more vulnerable groups such as children, persons with disabilities and others that have historically been disadvantaged or are at risk of exclusion, and to situations which are characterised by asymmetries of power or information, such as between employers and workers, or between businesses and consumers.
- Acknowledge that, while bringing substantial benefits to individuals and society, AI systems also pose certain risks and may have a negative impact, including impacts which may be difficult to anticipate, identify or measure (e.g. on democracy, the rule of law and distributive justice, or on the human mind itself.) Adopt adequate measures to mitigate these risks when appropriate, and proportionately to the magnitude of the risk.

Ethics Guidelines for Trustworthy AI

Requirements - Key guidance:

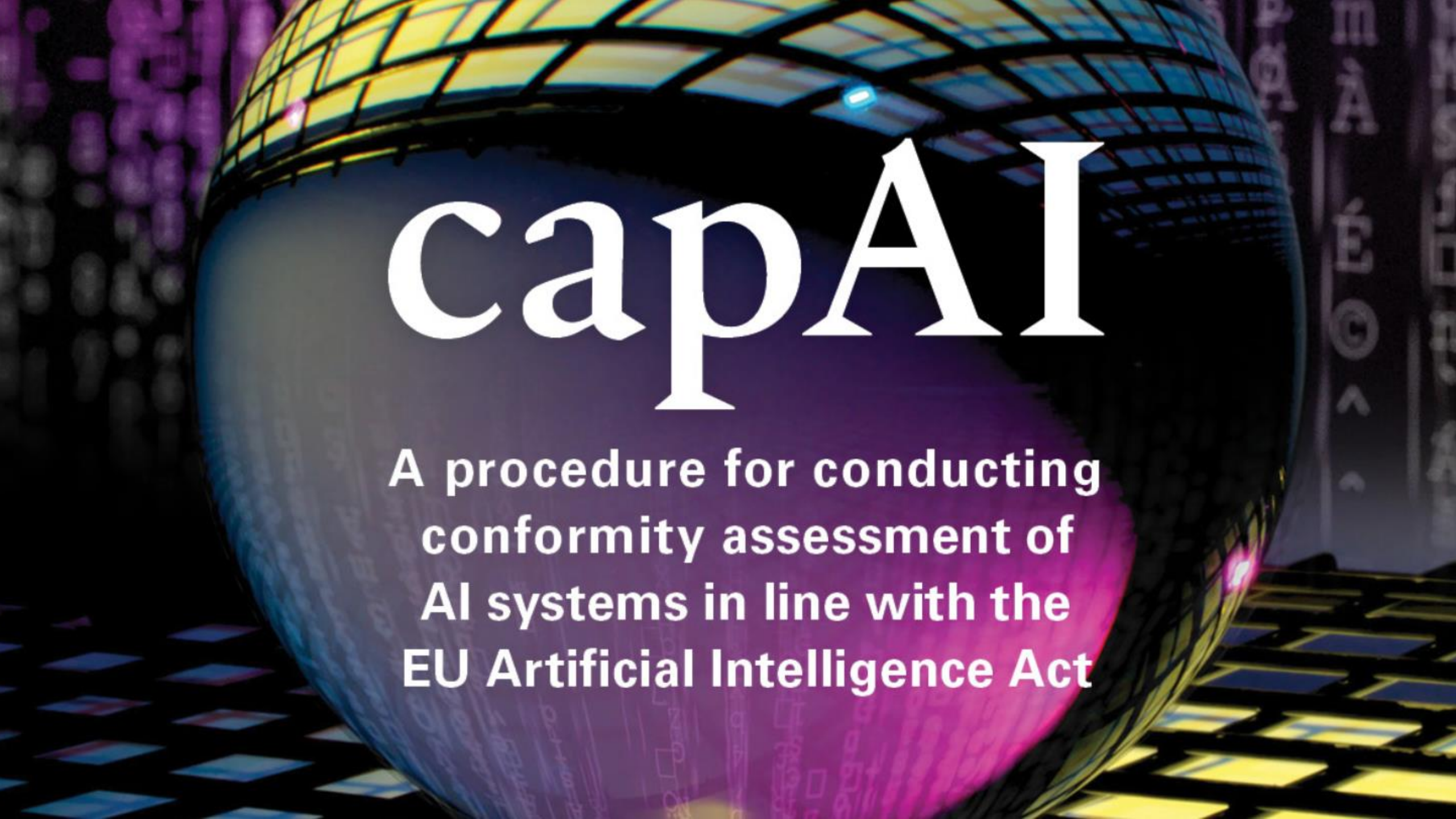
- Ensure that the development, deployment and use of AI systems meets the seven key requirements for Trustworthy AI: (1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, non-discrimination and fairness, (6) environmental and societal well-being and (7) accountability.
- Consider technical and non-technical methods to ensure the implementation of those requirements.
- Foster research and innovation to help assess AI systems and to further the achievement of the requirements; disseminate results and open questions to the wider public, and systematically train a new generation of experts in AI ethics.
- Communicate, in a clear and proactive manner, information to stakeholders about the AI system's capabilities and limitations, enabling realistic expectation setting, and about the manner in which the requirements are implemented. Be transparent about the fact that they are dealing with an AI system.
- Facilitate the traceability and auditability of AI systems, particularly in critical contexts or situations.
- Involve stakeholders throughout the AI system's life cycle. Foster training and education so that all stakeholders are aware of and trained in Trustworthy AI.
- Be mindful that there might be fundamental tensions between different principles and requirements. Continuously identify, evaluate, document and communicate these trade-offs and their solutions.

Ethics Guidelines for Trustworthy AI

Assessment list - Key guidance:

- Adopt a Trustworthy AI assessment list when developing, deploying or using AI systems, and adapt it to the specific use case in which the system is being applied.
- Keep in mind that such an assessment list will never be exhaustive. Ensuring Trustworthy AI is not about ticking boxes, but about continuously identifying and implementing requirements, evaluating solutions, ensuring improved outcomes throughout the AI system's lifecycle, and involving stakeholders in this.


European Commission
GENDER & INTERSECTIONAL BIAS IN ARTIFICIAL INTELLIGENCE
EUROPE FIT FOR THE DIGITAL AGE
@EUScinceInnov
#DigitalEU #UnionOfEquality
September 2020
Digital transformation and artificial intelligence (AI) are an integral part of the 4th Industrial Revolution, transforming our jobs and lives. But, while AI could well be a key driver for innovation, there are challenges and risks that cannot be ignored. This is why the Commission's White Paper on AI stresses the importance of building an AI 'ecosystem of trust' with 'rules that put people at the centre', as President von der Leyen stated in her 2020 State of the Union address.
DID YOU KNOW THAT:
Facial recognition systems perform better on men's faces than on women's, and on lighter skin than darker skin? Error rates vary from 35% for darker-skinned women, to 12% for darker-skinned men, 7% for lighter-skinned women, and less than 1% for lighter-skinned men. So, these systems need to be checked for bias and the people operating them trained accordingly[1].
Virtual assistants (e.g. chatbots) are often subjected to sexual harassment? Virtual assistants, e.g. Siri and Alexa, are usually programmed to respond to harassment with flirty, apologetic or deflecting responses. Research suggests that such responses perpetuate the stereotype of subservient women in service roles and may promote a culture of violence against women by presenting indirect ambiguity as a valid response to harassment. To address this issue, some companies have started developing software that is less tolerant of abuse[2].
Women are more likely to feel unwell when using virtual reality (VR), a technique which can be enhanced by AI? The symptoms experienced include i) pallor, ii) sweating, iii) increased heart rate, iv) drowsiness, v) disorientation and vi) general discomfort. So, it is important to test VR technologies on women, as well as men, and to promote gender balance in teams developing and designing VR and AI applications[3].


capAI
A procedure for conducting conformity assessment of AI systems in line with the EU Artificial Intelligence Act

Project results pack

- COSMIC DANCE used machine learning to understand how stars form and evolve.
- COG-TOM project developed novel data analytical methods called cognitive tomography to better understand the human mind.
- CoRob-X built cooperative robotic planetary rovers with environmental perception and control systems for autonomous decision-making ability in extreme environments.
- DeepCube analysed big data from the Copernicus Earth Observation Programme to improve our understanding of environmental challenges.
- DrugComb developed mathematical and computational tools to identify drug combinations for personalised cancer treatments.
- F-IMAGE employed AI to classify and analyse seismic signals and understand the behaviour of faults during earthquakes.
- HYPERION created a framework to assess the impacts of climate change and atmospheric composition on cultural heritage sites.
- MOUSSE exploited the similarities between languages to create a large network of language-independent semantic representations of sentences and tackle the challenge of AI hallucinations.
- NNNPDF explored the inner life of the proton and its structure using machine learning techniques.
- RD-ADVANCE used regression discontinuity design to better quantify the causal impact of different types of policy measures.
- STOP applied AI to data collection and analysis to create a platform for tailored nutrition for obesity sufferers.
- TAXINOMISIS used machine learning and neural networks to build a risk stratification platform for patients with carotid artery disease.
- TopMechMat created new materials with specialised properties, using AI to optimise their design.

FUNDAMENTAL RESEARCH

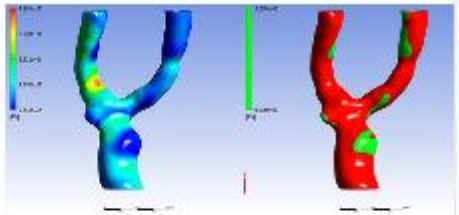
Project breaks new grounds in AI to create 'DNA of language'

25 August 2023

CLIMATE CHANGE AND ENVIRONMENT

AI helps preserve European cultural heritage

25 August 2023



HEALTH

AI in personalised medicine: Better risk identification, prevention and treatment of carotid artery disease

25 August 2023

HEALTH

Decoding the mind with the help of AI

25 August 2023

FUNDAMENTAL RESEARCH

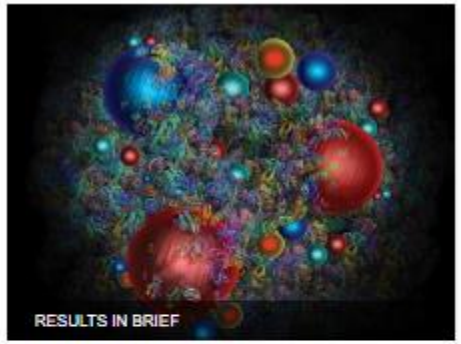
Machine learning improves the estimation of the causal effects in regression discontinuity designs (RDD)

25 August 2023

SPACE

Machine learning illuminates the celestial stage: spotting new objects amid the cosmic crowd

25 August 2023



FUNDAMENTAL RESEARCH

The enigmatic world of protons unveiled by AI

25 August 2023

SECURITY

AI could help anticipate catastrophic effects of future earthquakes

25 August 2023

HEALTH

AI-based platform offers personalised support to people with obesity

25 August 2023

HEALTH

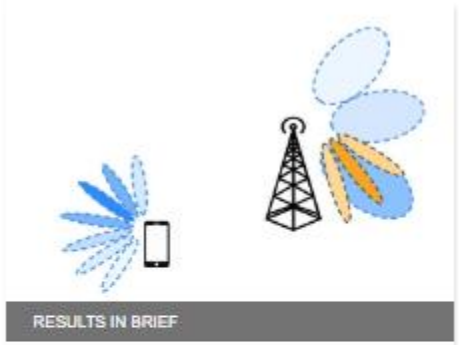
Unlocking the future of personalised cancer treatment with the help of AI

25 August 2023

SPACE

Mission accomplished: AI gives autonomy to space robots

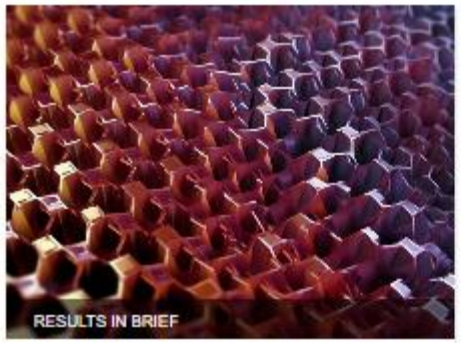
25 August 2023



DIGITAL ECONOMY

AI: helping wireless networks help themselves

25 August 2023



FUNDAMENTAL RESEARCH

Pioneering design platform reveals novel topological structures

25 August 2023

DIGITAL ECONOMY

Digital curation helps unveil historical truth about the Holocaust buried in archives

20 October 2022

SPACE

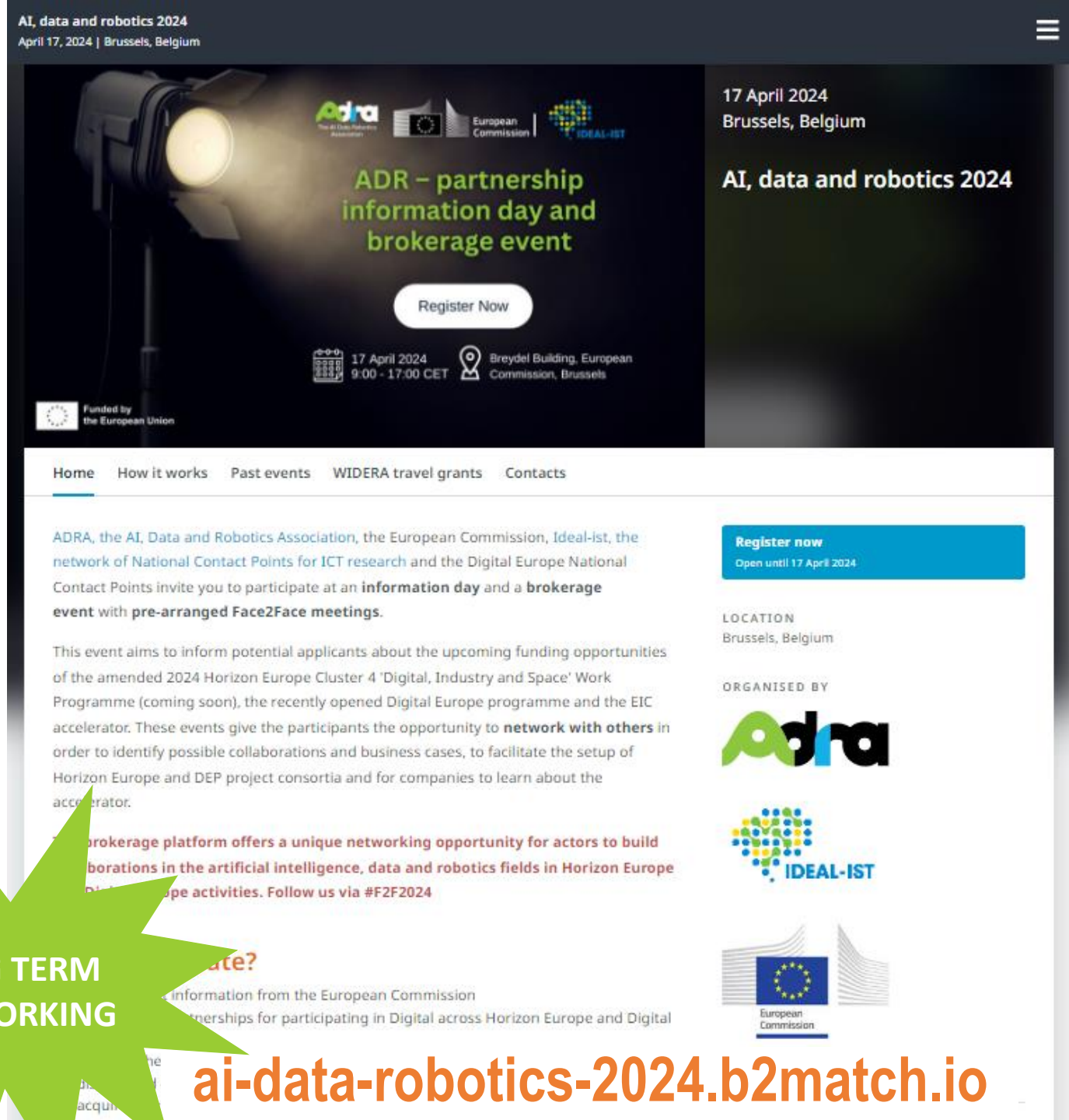
Infusing AI into Earth Observation: ushering in a new era in climate response

25 August 2023

https://cordis.europa.eu/article/id/446030-artificial-intelligence-expanding-scientific-boundaries-and-enhancing-innovation

VIRTUAL AGORA

- Register
- Create a profile
- Share cooperation offers
  - Project ideas
  - Expertise
  - Requests
- Find & contact partners



LONG TERM NETWORKING

IDEAL-IST

A service offered by:
the network of ICT
National Contact Points

ai-data-robotics-2024.b2match.io

CONNECT | DIGITIZE | TRANSFORM

Edina Nemeth

Programme Committee delegate, National Contact Point
Horizon Europe, Digital, Industry & Space (Cluster 4)
Horizon Europe, Cybersecurity (Cluster 3)
European Innovation Council – Pathfinder, Transition (Pillar 3)

National Research, Development and Innovation Office

+36-70-221-0387 - edina.nemeth@ist.hu

Ideal-ist – the network of National Contact Points
www.ideal-ist.eu