The European Union Artificial Intelligence Act

Major Harms and Emerging Risks Radar

August 2024

For more information on this publication, visit https://www.ai-and-partners.com/.

About AI & Partners

'AI That You Can Trust' - Your trusted advisor for EU AI Act Compliance. Unlock the full potential of artificial intelligence while ensuring compliance with the EU AI Act by partnering with AI & Partners, a leading professional services firm. We specialise in providing comprehensive and tailored software solutions for companies subject to the EU AI Act, guiding them through the intricacies of regulatory requirements and enabling responsible and accountable AI practices. At AI & Partners, we understand the challenges and opportunities that the EU AI Act presents for organisations leveraging AI technologies. Our team of seasoned experts combines in-depth knowledge of AI systems, regulatory frameworks, and industry specific requirements to deliver strategic guidance and practical solutions that align with your business objectives.

To find out how we can help you, email contact@ai-and-partners.com or visit https://www.ai-and-partners.com.

Business Integrity

AI & Partners defends and extends the digital rights of users at risk around the world. By combining direct technical support, comprehensive policy engagement, global advocacy, grassroots professional services, regulatory interventions, and participating in industry groups such as AI Commons, we fight for fundamental rights in the artificial intelligence age.

AI & Partners' publications do not necessarily reflect the opinions of its clients, partners and/or stakeholders.

This document is published by AI & Partners as part of its ongoing contributions to EU AI Act preparations, insight areas or interactions. The findings, interpretations and conclusions expressed herein are a result of a collaborative process facilitated and endorsed by AI & Partners but whose results do not necessarily represent the views of the AI & Partners, nor the entirety of its Partners or other stakeholders.

Contents

AI & Partners

Amsterdam - London - Singapore

We are pleased to present the 2024 version of our Emerging Risk Radar.

Emerging Risks are risks which may newly develop or which already exist and are continuously evolving. They are characterised by a high degree of uncertainty in terms of impacts and likelihood, according to the EU AI Act, and may have a substantial potential impact on an enterprise's activity.

The Radar is a summary of emerging risks and associated major impacts, protection types, and risk levels that could affect enterprises across industries affected by and/or subject to the EU AI Act over the next five years and beyond. Risks are classified as unacceptable, high, specific transparency, and minimal based on their EU AI Act applicability. Both the list of risks and the type of impact are based on the expert opinion of specialists and market practitioners at AI & Partners leveraging over 3.5 years of experience.

The major impacts have been constructed in line with their reference under Recital 5.

All risks have been assessed and created as part of brainstorming sessions together with extensive market research.

We hope you find the report useful and welcome your comments and feedback.

Executive Summary | Risk Pyramid | Major Harm Descriptions | Protection Descriptions | Risk Level Descriptions | Emerging Risk Descriptions
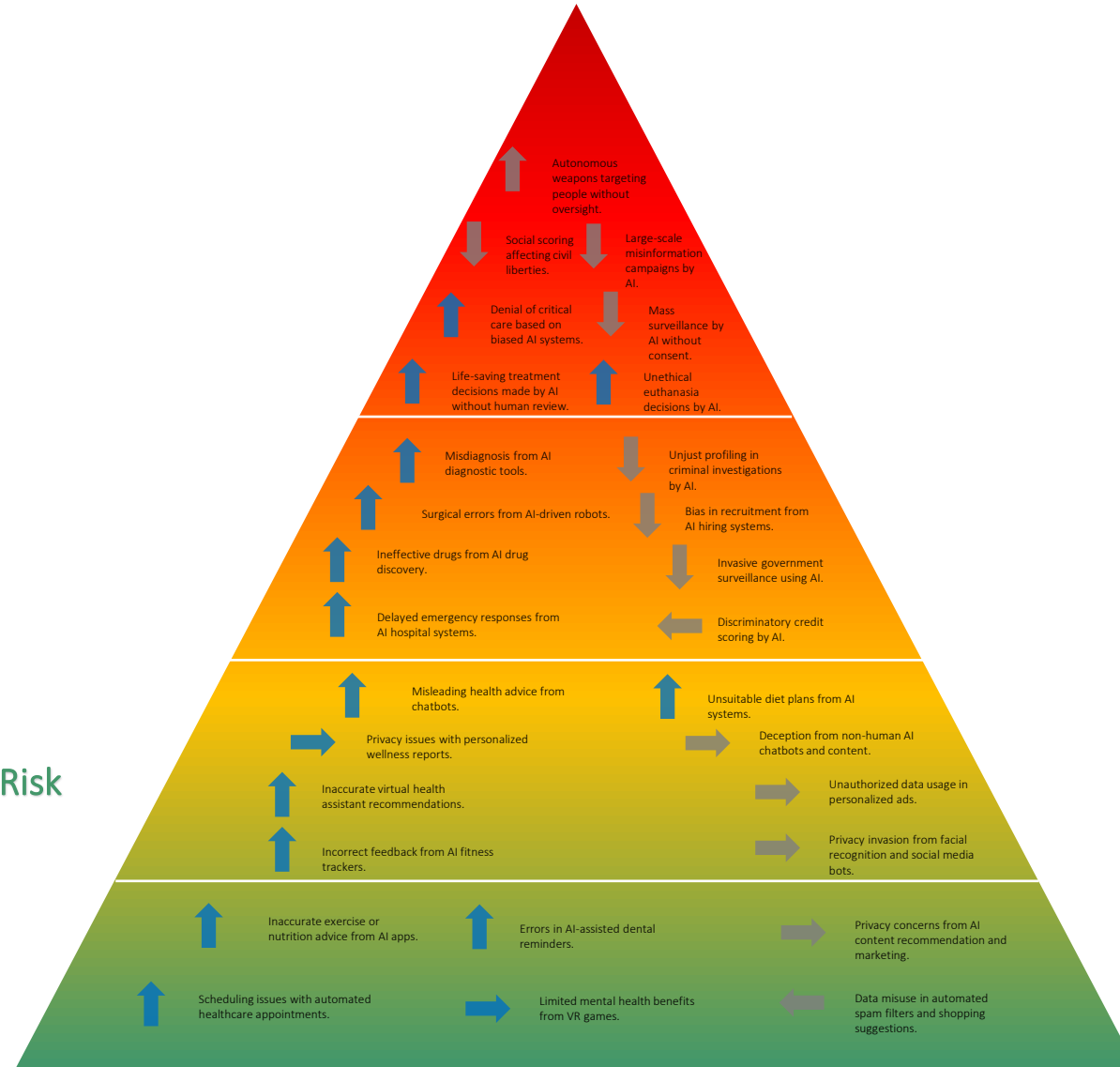
Risk Level

Unacceptable Risk

High Risk

Specific Transparency Risk

Minimal Risk

Type

- Health & Safety
- Fundamental Rights

Harm

- Physical
- Psychological
- Societal
- Economic

Pyramid contents (top to bottom):

Autonomous weapons targeting people without oversight.

Social scoring affecting civil liberties.

Large-scale misinformation campaigns by AI.

Denial of critical care based on biased AI systems.

Mass surveillance by AI without consent.

Life-saving treatment decisions made by AI without human review.

Unethical euthanasia decisions by AI.

Misdiagnosis from AI diagnostic tools.

Unjust profiling in criminal investigations by AI.

Surgical errors from AI-driven robots.

Bias in recruitment from AI hiring systems.

Ineffective drugs from AI drug discovery.

Invasive government surveillance using AI.

Delayed emergency responses from AI hospital systems.

Discriminatory credit scoring by AI.

Misleading health advice from chatbots.

Unsuitable diet plans from AI systems.

Privacy issues with personalized wellness reports.

Deception from non-human AI chatbots and content.

Inaccurate virtual health assistant recommendations.

Unauthorized data usage in personalized ads.

Incorrect feedback from AI fitness trackers.

Privacy invasion from facial recognition and social media bots.

Inaccurate exercise or nutrition advice from AI apps.

Errors in AI-assisted dental reminders.

Privacy concerns from AI content recommendation and marketing.

Scheduling issues with automated healthcare appointments.

Limited mental health benefits from VR games.

Data misuse in automated spam filters and shopping suggestions.

Executive Summary | **Risk Pyramid** | Major Harm Descriptions | Protection Descriptions | Risk Level Descriptions | Emerging Risk Descriptions

Harm caused to public interests

Physical

Physical impact refers to any direct or indirect effect that an AI system can have on the physical well-being of individuals or the physical environment. This includes:

- **Bodily Harm**: Direct Harm: Injuries or fatalities caused directly by the malfunction or misuse of an AI system, such as an autonomous vehicle causing an accident; and Indirect Harm: Situations where the AI system's decisions or actions lead to conditions that result in physical harm, like an AI managing industrial machinery that fails, leading to workplace accidents.
- **Damage to Property**: Infrastructure Disruption: AI systems controlling critical infrastructure, such as power grids or water supply systems, could cause significant physical damage if they malfunction or are compromised; and Property Damage: AI systems used in home automation or security could potentially cause damage to personal property if they fail or are hacked.
- **Environmental Impact**: Pollution and Contamination: AI systems used in industrial processes could lead to environmental harm if they malfunction, resulting in pollution or contamination; and Resource Depletion: AI systems that manage natural resources could cause physical environmental damage if not properly regulated, leading to over-extraction or habitat destruction.

Psychological

Psychological impact refers to the effects that the use of AI systems can have on the mental and emotional well-being of individuals. This includes:

- **Emotional Well-being**: AI systems, especially those involving biometric data and emotion recognition, can influence an individual's emotional state. For instance, an AI system that misinterprets emotions could lead to stress or anxiety.
- **Behavioural Influence**: AI systems designed to predict or influence behaviour can have significant psychological effects. For example, systems used in advertising or social media can affect user behaviour and mental health by manipulating content exposure.
- **Privacy Concerns**: The use of AI systems that process sensitive personal data, including biometric data, can lead to psychological stress due to privacy invasion concerns. Individuals may feel a loss of control over their personal information, leading to anxiety and distrust.

Societal

Societal impact pertains to the broader effects on society as a whole. This includes:

- **Changes in Social Structures**: AI systems can alter community dynamics and cultural norms. For example, widespread use of AI in the workplace could change employment patterns and social interactions.
- **Public Trust**: The deployment of AI systems can affect public trust in technology and institutions. Misuse or failures of AI systems can lead to a loss of trust in these technologies.
- **Discrimination and Inequality**: AI systems can perpetuate or exacerbate existing social inequalities. For instance, biased algorithms can lead to discriminatory practices in hiring, lending, or law enforcement.

Economic

Economic impact involves the financial and economic consequences of AI systems. This includes:

- **Job Displacement**: AI systems can automate tasks previously performed by humans, leading to job losses in certain sectors. However, they can also create new job opportunities in other areas.
- **Market Dynamics**: AI can change market dynamics by optimizing operations and resource allocation, potentially leading to increased productivity and economic growth.
- **Economic Inequality**: The benefits of AI may not be evenly distributed, potentially leading to increased economic inequality. Companies and individuals with access to advanced AI technologies may gain a competitive advantage over those without.

Executive Summary | Risk Pyramid | **Major Harm Descriptions** | Protection Descriptions | Risk Level Descriptions | Emerging Risk Descriptions

AI & Partners

Amsterdam - London - Singapore

Health & Safety

This is not explicitly defined in the EU AI Act under Article 3. However, the Act frequently references "health and safety" in the context of high-risk AI systems and their potential impacts. Here are the key points:

- **High-Risk AI Systems**: The Act categorizes certain AI systems as high-risk based on their potential to harm health and safety or adversely impact fundamental rights.
- **Conformity Assessment**: The Act mandates conformity assessments for high-risk AI systems to ensure they meet safety requirements, thereby protecting health and safety.
- **Amendments and Standards**: The Act includes provisions for adopting standards and amendments to maintain the level of protection for health and safety.
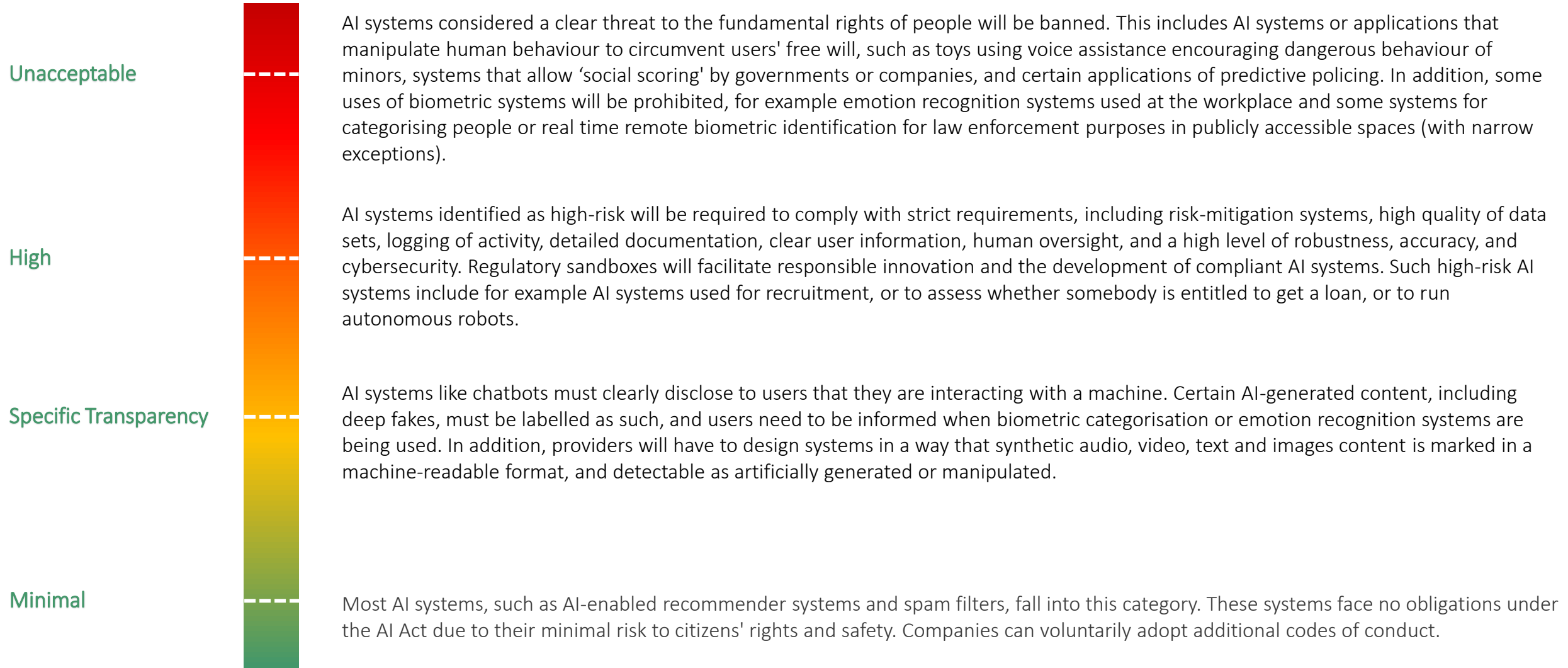
Fundamental Rights

This is also not explicitly defined within the EU AI Act under Article 3. However, the Act frequently refers to the protection of fundamental rights, emphasizing their importance in various contexts. Here are the key points:

- **Protection of Fundamental Rights**: The Act aims to ensure a high level of protection for fundamental rights, including health, safety, and the rights enshrined in the Charter of Fundamental Rights of the European Union.
- **Obligations for High-Risk AI Systems**: Providers and deployers of high-risk AI systems must conduct a fundamental rights impact assessment to evaluate the potential risks and impacts on fundamental rights before deploying such systems.

Executive Summary | Risk Pyramid | Major Harm Descriptions | **Protection Descriptions** | Risk Level Descriptions | Emerging Risk Descriptions

European Union Artificial Intelligence Act ● Major Harms and Emerging Risks Radar ● AI & Partners ● https://www.ai-and-partners.com/   7

Risk-based approach to AI system regulation

Unacceptable

AI systems considered a clear threat to the fundamental rights of people will be banned. This includes AI systems or applications that manipulate human behaviour to circumvent users' free will, such as toys using voice assistance encouraging dangerous behaviour of minors, systems that allow 'social scoring' by governments or companies, and certain applications of predictive policing. In addition, some uses of biometric systems will be prohibited, for example emotion recognition systems used at the workplace and some systems for categorising people or real time remote biometric identification for law enforcement purposes in publicly accessible spaces (with narrow exceptions).

High

AI systems identified as high-risk will be required to comply with strict requirements, including risk-mitigation systems, high quality of data sets, logging of activity, detailed documentation, clear user information, human oversight, and a high level of robustness, accuracy, and cybersecurity. Regulatory sandboxes will facilitate responsible innovation and the development of compliant AI systems. Such high-risk AI systems include for example AI systems used for recruitment, or to assess whether somebody is entitled to get a loan, or to run autonomous robots.

Specific Transparency

AI systems like chatbots must clearly disclose to users that they are interacting with a machine. Certain AI-generated content, including deep fakes, must be labelled as such, and users need to be informed when biometric categorisation or emotion recognition systems are being used. In addition, providers will have to design systems in a way that synthetic audio, video, text and images content is marked in a machine-readable format, and detectable as artificially generated or manipulated.

Minimal

Most AI systems, such as AI-enabled recommender systems and spam filters, fall into this category. These systems face no obligations under the AI Act due to their minimal risk to citizens' rights and safety. Companies can voluntarily adopt additional codes of conduct.

Risks preventing the uptake of human-centric and trustworthy AI

Risk	Risk Level	Associated Harm	Associated Protection	Description
Inaccurate exercise or nutrition advice from AI apps.	Minimal	Physical	Health & Safety	Minimal risk as recommendations are non-invasive and personalized for general fitness, with no critical health decisions involved.
Scheduling issues with automated healthcare appointments.	Minimal	Physical	Health & Safety	Automates appointment reminders, reducing the risk of missed visits with no impact on treatment decisions.
Errors in AI-assisted dental reminders.	Minimal	Physical	Health & Safety	Reminds patients of upcoming dental visits, posing minimal risk by enhancing healthcare access without influencing care quality.
Limited mental health benefits from VR games.	Minimal	Psychological	Health & Safety	Minimal risk as these provide non-invasive relaxation, offering stress relief through immersive experiences without medical intervention.
Privacy concerns from AI content recommendation and marketing.	Minimal	Psychological	Fundamental Rights	Minimal risk as AI simply enhances user experience by suggesting content, with little impact on rights or freedoms.
Data misuse in automated spam filters and shopping suggestions.	Minimal	Economic	Fundamental Rights	Minimal risk, where AI assists in personalizing shopping experiences, with negligible impact on fundamental rights.

Risks preventing the uptake of human-centric and trustworthy AI

Risk	Risk Level	Associated Harm	Associated Protection	Description
Misleading health advice from chatbots.	Specific Transparency	Physical	Health & Safety	Requires clear disclosure to users that advice comes from AI, as it could be mistaken for expert medical guidance.
Privacy issues with personalized wellness reports..	Specific Transparency	Psychological	Health & Safety	Needs transparency in data sources and algorithms to avoid misleading interpretations of wellness assessments.
Inaccurate virtual health assistant recommendations.	Specific Transparency	Physical	Health & Safety	Potential for users to over-rely on AI without understanding its limitations, necessitating clear communication of AI's role.
Incorrect feedback from AI fitness trackers.	Specific Transparency	Physical	Health & Safety	Feedback on physical activity requires clear explanation of how data is analyzed to avoid confusion about fitness guidance.
Unsuitable diet plans from AI systems.	Specific Transparency	Physical	Health & Safety	Transparency in how diet plans are generated is crucial to prevent misuse or misunderstanding of personalized nutritional advice.
Deception from non-human AI chatbots and content.	Specific Transparency	Psychological	Fundamental Rights	Transparency is essential to differentiate between AI-generated and human-written content, avoiding misinformation.
Unauthorized data usage in personalized ads.	Specific Transparency	Psychological	Fundamental Rights	Transparency about data collection and usage is necessary to maintain trust and protect user privacy.
Privacy invasion from facial recognition and social media bots.	Specific Transparency	Psychological	Fundamental Rights	Transparency required to prevent misleading interactions, ensuring users are aware they're engaging with AI.

Executive Summary | Risk Pyramid | Major Harm Descriptions | Protection Descriptions | Risk Level Descriptions | **Emerging Risk Descriptions**

Risks preventing the uptake of human-centric and trustworthy AI

Risk	Risk Level	Associated Harm	Associated Protection	Description
Misdiagnosis from AI diagnostic tools.	High	Physical	Health & Safety	High risk due to potential misdiagnosis or false positives, impacting critical healthcare decisions and patient outcomes.
Surgical errors from AI-driven robots.	High	Physical	Health & Safety	High risk as AI assists in surgeries where precision is critical, with potential life-threatening consequences if errors occur.
Ineffective drugs from AI drug discovery.	High	Physical	Health & Safety	High risk associated with creating new drugs, as errors in AI analysis could lead to ineffective or dangerous medications.
Delayed emergency responses from AI hospital systems.	High	Physical	Health & Safety	High risk if AI systems fail in emergency settings, potentially delaying life-saving interventions or misallocating resources.
Unjust profiling in criminal investigations by AI.	High	Societal	Fundamental Rights	High risk due to potential for bias and infringement on civil liberties, impacting fair treatment in justice systems.
Bias in recruitment from AI hiring systems.	High	Societal	Fundamental Rights	High risk as biased algorithms can lead to discrimination in hiring, affecting equal employment opportunities.
Invasive government surveillance using AI.	High	Societal	Fundamental Rights	High risk of infringing on privacy rights, with potential for misuse in monitoring and controlling populations.
Discriminatory credit scoring by AI.	High	Economic	Fundamental Rights	High risk of infringing on privacy rights, with potential for misuse in monitoring and controlling populations.

Risks preventing the uptake of human-centric and trustworthy AI

Risk	Risk Level	Associated Harm	Associated Protection	Description
Denial of critical care based on biased AI systems.	Unacceptable	Physical	Health & Safety	Unacceptable risk involving discrimination in healthcare access, violating fundamental rights to equitable treatment.
Life-saving treatment decisions made by AI without human review.	Unacceptable	Physical	Health & Safety	Unacceptable risk as it removes human judgment from critical, life-altering decisions, increasing the potential for fatal errors.
Unethical euthanasia decisions by AI.	Unacceptable	Physical	Health & Safety	Unacceptable risk involving irreversible decisions about life and death, requiring absolute human control and ethical consideration.
Mass surveillance by AI without consent.	Unacceptable	Physical	Fundamental Rights	Unacceptable risk due to the severe infringement on privacy and civil liberties, leading to potential state overreach.
Large-scale misinformation campaigns by AI.	Unacceptable	Societal	Fundamental Rights	Unacceptable risk due to the potential to distort public opinion and erode trust in information sources, impacting democratic processes.
Social scoring affecting civil liberties.	Unacceptable	Societal	Fundamental Rights	Unacceptable risk involving the reduction of individuals' rights and freedoms based on AI assessments, leading to social inequality.
Autonomous weapons targeting people without oversight.	Unacceptable	Societal	Fundamental Rights	Unacceptable risk, as it removes human judgment from lethal decisions, raising severe ethical and legal concerns.

— Thank you!

Amsterdam - London - Singapore

Email
contact@ai-and-partners.com

Phone
+44(0)7535 994 132

Website
https://www.ai-and-partners.com/

Social Media
LinkedIn: https://www.linkedin.com/company/ai-&-partners/
Twitter: https://twitter.com/AI_and_Partners

This Presentation may contain information, text, data, graphics, photographs, videos, sound recordings, illustrations, artwork, names, logos, trade marks, service marks, and information about us, our lines of services, and general information may be provided in the form of documents, podcasts or via an RSS feed ("the Information").

Except where it is otherwise expressly stated, the Information is not intended to, nor does it, constitute legal, accounting, business, financial, tax or other professional advice or services. The Information is provided on an information basis only and should not be relied upon. If you need advice or services on a specific matter, please contact us using the contact details for the relevant consultant or fee earner found on the Presentation.

The Presentation and Information is provided "AS IS" and on an "AS AVAILABLE" basis and we do not guarantee the accuracy, timeliness, completeness, performance or fitness for a particular purpose of the Presentation or any of the Information. We have tried to ensure that all Information provided on the Presentation is correct at the time of publication. No responsibility is accepted by or on behalf of us for any errors, omissions, or inaccurate information on the Presentation. Further, we do not warrant that the Presentation or any of the Information will be uninterrupted or error-free or that any defects will be corrected.

Although we attempt to ensure that the Information contained in this Presentation is accurate and up-to-date, we accept no liability for the results of any action taken on the basis of the Information it contains and all implied warranties, including, but not limited to, the implied warranties of satisfactory quality, fitness for a particular purpose, non-infringement, compatibility, security, and accuracy are excluded from these Terms to the extent that they may be excluded as a matter of law.

In no event will we be liable for any loss, including, without limitation, indirect or consequential loss, or any damages arising from loss of use, data or profits, whether in contract, tort or otherwise, arising out of, or in connection with the use of this Presentation or any of the Information.