AI & Partners

Amsterdam - London - Singapore

AI System Database Report

Forward-looking insights for businesses

February 2024

Introduction

- A European Union ("EU") database of high-risk Artificial intelligence ("AI") systems (the "Database") is a key tool for *transparency* in the upcoming European Union AI Act (the "EU AI Act").

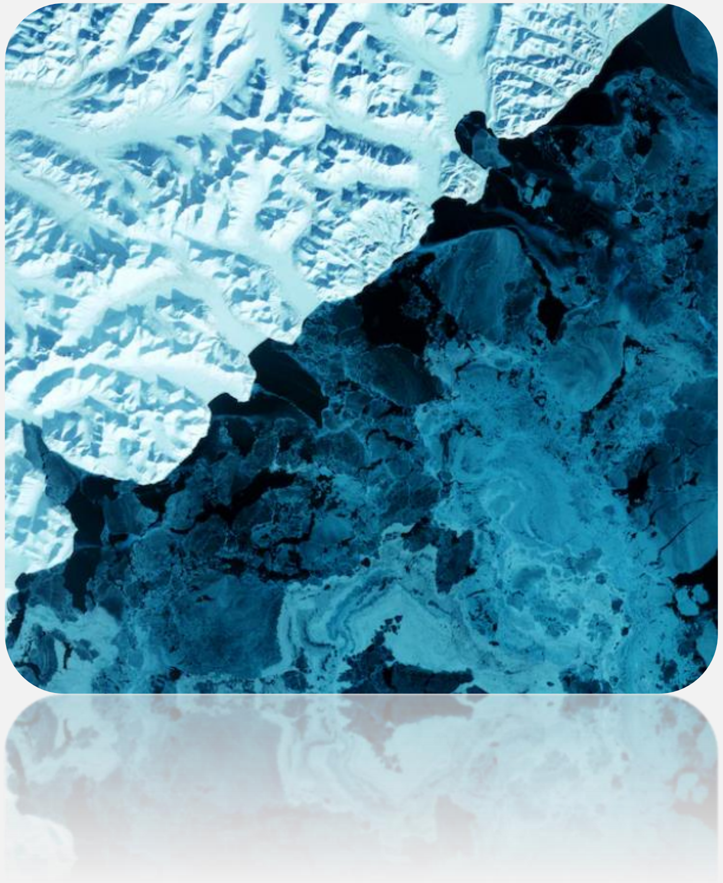
- The Database is intended to be freely and publicly accessible, easily understandable and machine-readable. Moreover, the Database aims to be user-friendly and easily navigable, with search functionalities at minimum allowing the general public to search the database for specific high-risk systems, locations, categories of risk under Annex IV and keywords.

- Ahead of the EU AI Act's scheduled entry into force in 2024, this document looks at top ten *potential* product categories, hazards and corrective actions for deployers and/or developers of high-risk AI systems, using the United Kingdom's ("UK") '**Product Safety Database Report from 2022 to 2023**', as the classification of risk under the EU AI Act is grounded in the intended purpose of the AI system, aligning with existing EU product safety legislation.

- **Recital 69** of the EU AI Act requires providers of high-risk AI systems to register their high-risk AI system and foundation models in an EU database.

*All data sourced from the European Commission (2023)

Product Categories



Facial Recognition Systems for Law Enforcement

Fits into the high-risk category, especially if used in real-time remote biometric identification in publicly accessible spaces.

Predictive Policing Software

Falls under the unacceptable risk category, as individual predictive policing is banned.

Emotion Recognition Software for Non-Medical or Non-Safety Purposes

Considered unacceptable risk, as emotion recognition in workplaces and educational institutions is prohibited except for medical or safety reasons.

Social Scoring Systems

Falls into the unacceptable risk category, as the use of AI for public and private social scoring purposes is banned.

Vulnerability Exploitation Software

Classified as an unacceptable risk, as the exploitation of vulnerabilities of individuals using AI is prohibited.

Biometric Categorization Systems Based on Sensitive Data

Falls into the unacceptable risk category, as biometric categorization of individuals based on sensitive data is banned unless used to identify victims.

Untargeted Scraping Tools for Facial Images

Considered an unacceptable risk, as the untargeted scraping of the internet or CCTV for facial images to build or expand databases is banned.

AI Systems for Individualized Social and Predictive Profiling

Falls under the high-risk category, especially if used in ways that impact fundamental rights or safety.

AI Systems Used in Sector-Specific Products Subject to Union Legislation

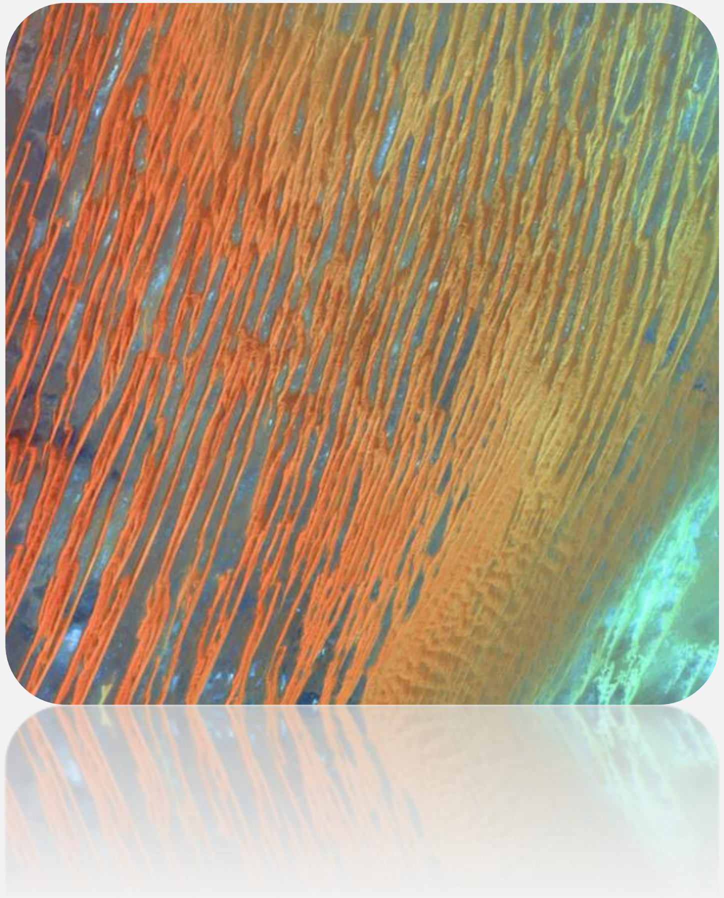
High-risk AI systems encompassing safety components of products covered by sector-specific Union legislation, remaining high-risk when subjected to third-party conformity assessment.

Biometric Data Filtering Tools in Law Enforcement

Falls into the unacceptable risk category but note that filtering datasets based on biometric data in law enforcement is still possible, though limited.

*All data sourced from the European Commission (2023)

Hazards



Privacy Violations

Hazard associated with the collection and use of facial recognition data, emotion data, and biometric information without proper consent or in violation of privacy laws.

Algorithmic Bias

Hazard related to the potential biases in predictive policing software, social scoring systems, and biometric categorization systems, leading to unfair or discriminatory outcomes.

Security Vulnerabilities

Hazard associated with vulnerabilities in facial recognition systems, predictive policing software, and other AI systems, leading to security breaches and unauthorized access.

Misuse of Emotion Recognition

Hazard related to the misuse of emotion recognition software for non-medical or non-safety purposes, potentially leading to emotional manipulation or privacy infringements.

Social Manipulation

Hazard associated with the use of social scoring systems for public and private purposes, leading to potential manipulation of individuals and social discrimination.

Exploitation of Vulnerabilities

Hazard related to the use of AI to exploit vulnerabilities in individuals, leading to potential harm, manipulation, or unauthorized access to sensitive information.

Biometric Data Misuse

Hazard associated with the misuse of biometric categorization systems based on sensitive data, potentially leading to discrimination, profiling, or privacy violations.

Unintended Consequences of Predictive Policing

Hazard related to the unintended consequences of predictive policing software, such as reinforcing existing biases, infringing on civil liberties, or impacting marginalized communities disproportionately.

Unauthorized Data Scraping

Hazard associated with the use of untargeted scraping tools for facial images, leading to the unauthorized collection of sensitive information and potential misuse.

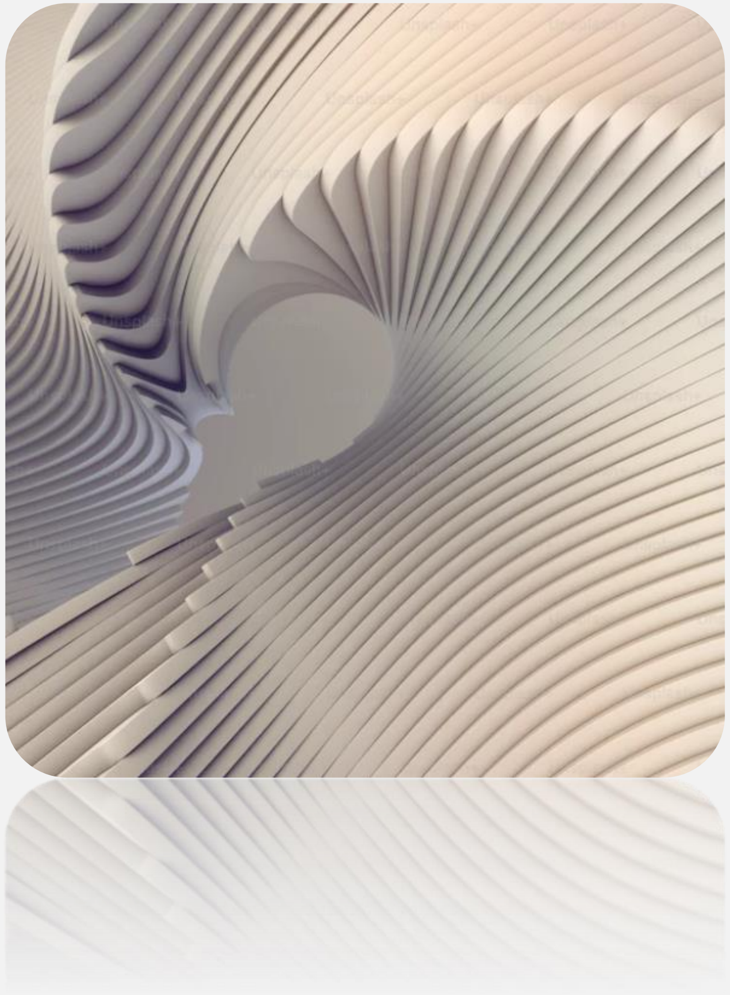
Lack of Transparency and Accountability

Hazard related to the lack of transparency in AI systems, especially in high-risk applications, leading to challenges in understanding decision-making processes and holding responsible parties accountable.

Introduction | Product Categories | **Hazards** | Corrective Actions

*All data sourced from the European Commission (2023)

Corrective Actions



Algorithmic Bias Mitigation

Implement measures to identify and mitigate biases in facial recognition systems, predictive policing software, and other AI products, ensuring fair and equitable outcomes.

Privacy Safeguards

Strengthen privacy controls in AI systems, especially in facial recognition and biometric categorization, to ensure compliance with regulations and protect individuals' privacy rights.

Security Patching and Regular Updates

Implement a rigorous schedule for security patching and updates to address vulnerabilities in AI systems, reducing the risk of security breaches and unauthorized access.

Transparency Measures

Enhance transparency in AI decision-making processes, providing explanations for predictions and actions taken by systems, especially in high-risk applications.

Ethical Guidelines Implementation

Incorporate and adhere to ethical guidelines in the development and deployment of AI systems, addressing concerns related to the impact on fundamental rights and societal values.

Compliance Audits

Conduct regular audits to ensure that AI products comply with relevant regulations and standards, particularly those related to social scoring, emotion recognition, and other high-risk applications.

User Education and Awareness

Provide users with information and education about the capabilities, limitations, and ethical considerations of AI products, fostering awareness and informed use.

Limiting Biometric Data Use

Restrict the use of biometric categorization systems based on sensitive data, ensuring compliance with regulations and preventing misuse.

Prohibited Functionality Checks

Implement checks within AI systems to ensure that prohibited functionalities, such as individual predictive policing and untargeted scraping, are not present or enabled.

Third-Party Conformity Assessment

Subject high-risk AI systems, especially those used in sector-specific products subject to Union legislation, to third-party conformity assessment to validate compliance with safety standards and regulations.

*All data sourced from the European Commission (2023)

Contact Details

AI
AI & Partners
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

AI

AI & Partners

Amsterdam - London - Singapore

Thank You!

Disclaimer

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.