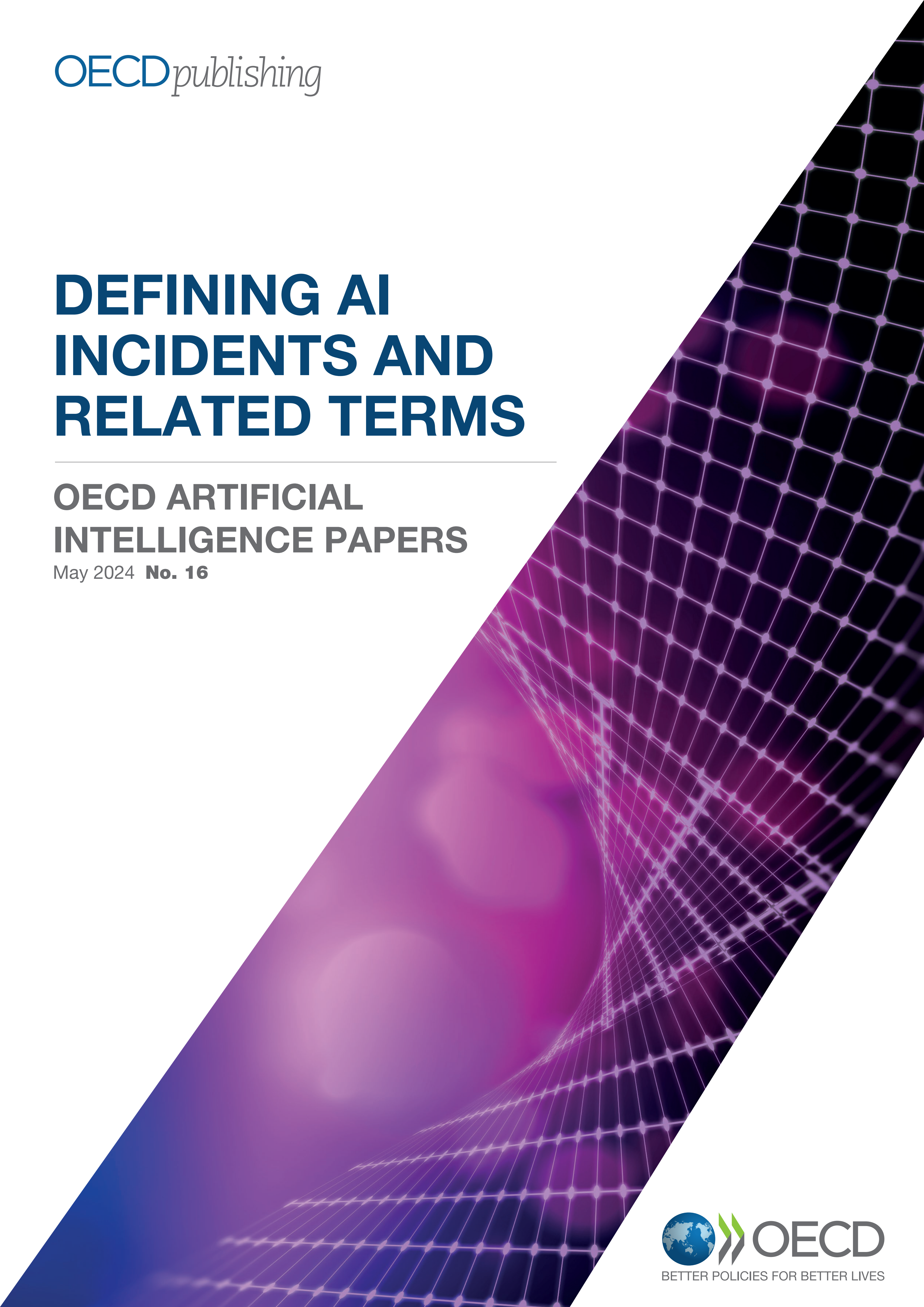
OECD*publishing*

DEFINING AI INCIDENTS AND RELATED TERMS

OECD ARTIFICIAL INTELLIGENCE PAPERS

May 2024 **No. 16**



OECD
BETTER POLICIES FOR BETTER LIVES

Foreword

This report provides a definition of AI incident and related terminology. These definitions aim to clarify what constitutes an AI incident, an AI hazard, and associated terminology without being overly prescriptive. This document will inform the development of a common AI incidents reporting framework and its application through the OECD AI Incidents Monitor (AIM).

This report and previous versions of it were discussed and reviewed by members of the former OECD.AI Expert Group on Classifying AI systems during a series of informal workshops between July and October 2022. It was also discussed by the OECD.AI Expert Group on AI Incidents at its March, April, June, August and November 2023 meetings and its February 2024 meeting. The OECD Working Party on Artificial Intelligence (AIGO) discussed this report and previous versions of it at its November 2022; April, July and November 2023; and February 2024 meetings.

The report was written by Karine Perset and Luis Aranda under the supervision of Audrey Plonk, Deputy Director of the OECD Science Technology and Innovation Directorate. The report also benefitted from the inputs of delegates for the OECD Working Party on Artificial Intelligence (AIGO), including the Civil Society Information Society Advisory Council (CSISAC) and Business at the OECD (BIAC). Bénédicte Rispal, Orsolya Dobe, John Tarver, Shellie Phillips and Andreia Furtado provided editorial support.

This paper was approved and declassified by written procedure by the OECD Digital Policy Committee (DPC) on 14 March 2024 and prepared for publication by the OECD Secretariat.

Note to Delegations:

This document is also available on O.N.E under the reference code:

DSTI/CDEP/AIGO(2023)10/FINAL

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Acknowledgements

This report is based on the work of the OECD.AI Expert Group on AI Incidents and the former OECD.AI Expert Group on Classifying AI systems. It was prepared under the aegis of the OECD Working Party on AI Governance (AIGO). The OECD.AI Expert Group on AI incidents was co-chaired by Irina Orssich (European Commission); Elham Tabassi (National Institute of Standards and Technology), Marko Grobelnik (Jozef Stefan Institute) and Mark Latonero (White House Office of Science and Technology Policy). Karine Perset and Luis Aranda (OECD Digital Economy Policy Division) led the report development and drafting. The authors thank Kevin Paeth (UL Research institute) for initial research, stocktaking and drafting.

Many experts provided invaluable feedback and suggestions, including Sean McGregor (Responsible AI Collaborative); Heather Frase (CSET); Irina Orssich, Tatjana Evas and Yordanka Ivanova (European Comission); Pam Dixon (World Privacy Forum); Kevin Paeth (UL Research institute); and Andrea Renda (CEPS); as well as of national delegations to the AIGO, including Roxane Sabourin (Canada); Ferenc Kasa (Hungary); Luis Ricardo Sánchez Hernández (Mexico); Rose Woolhouse (United Kingdom); Philipp Schulte (Germany); Ziv Katzir (Israel); David Turnbull (United States); and Ayaka Tano (Japan).

The paper benefitted significantly from the oral and written contributions of those associated with the OECD.AI Expert Group on AI Incidents, including Ahmet Yildiz (Ministry of Industry and Technology, Republic of Türkiye); Andrejs Vasiljevs (Tilde); Annalore Verhagen (OECD); Aurelie Jacquet (Standards Australia); Barry O'Brien (IBM); Carlos Ignacio Gutierrez (Future of Life Institute); Carlos Muñoz Ferrandis (BigScience); Catelijne Muller (ALLAI); Coran Darling (DLA Piper); Craig Shank (independent expert); Daniel Schwabe (Catholic University in Rio de Janeiro); Elham Tabassi (NIST); Eva Thelisson (AI Transparency Institute); Florian Ostmann (The Alan Turing Institute); Heather Frase (CSET); Ilya Meyzin (Dun & Bradstreet); Irina Orssich (European Commission); Jana Novohradska (Slovak Republic); Jessica Cussins (UC Berkeley); John McCarthy (Arup); Judith Peterka (Germany); Leonidas Aristodemou (OECD); Mark Latonero (White House Office of Science and Technology Policy); Marko Grobelnik (Jozef Stefan Institute); Mohammed Motiwala (United States); Nicolas Miailhe (The Future Society); Nicolas Moës (The Future Society); Nozha Boujemaa (IKEA); Pam Dixon (World Privacy Forum); Peter Addo (Agence Française de Développement); Peter Cihon (GitHub); Philip Dawson (Armila); Prateek Sibal (UNESCO); Raja Chatila (IEEE); Rayid Ghani (Carnegie Mellon University); Rebecca Anselmetti (United Kingdom); Sean McGregor (Responsible AI Collaborative); Sebastian Hallensleben (CEN-CENELEC); Tatjana Evas (European Commission); Thiago Guimarães Moraes (Brazil); Till Klein (AppliedAI); Yolanda Lannquist (The Future Society); and Yordanka Ivanova (European Commission).

The Secretariat would also like to thank stakeholder groups at the OECD for their input, including Pam Dixon (Civil Society Information Society Advisory – CSISAC); Nicole Primmer and Maylis Berviller (Business at OECD – BIAC); and Sarah Jameson and Aida Ponce (Trade Union Advisory Committee – TUAC).

Finally, the authors thank Bénédicte Rispal, Orsolya Dobe, John Tarver, Shellie Phillips, Andreia Furtado and Angela Gosmann for editorial support. The overall quality of this report benefitted significantly from their engagement.

Table of contents

FIGURES

TABLES

Abstract

As AI use grows, so do its benefits and risks. These risks can lead to actual harms, *AI incidents*, or potential dangers, *AI hazards*. Clear definitions are essential for managing and preventing these risks. This report proposes definitions for AI incidents and related terms. These definitions aim to foster international interoperability while providing flexibility for jurisdictions to determine the scope of AI incidents and hazards they wish to address.

Résumé

À mesure que l’utilisation de l’IA se développe, ses avantages et ses risques augmentent également. Ces risques peuvent entrainer des préjudices réels, incidents liés à l’IA, ou des dangers potentiels, dangers liés à l’IA. Des définitions claires sont essentielles pour la gestion et la prévention de ces risques. Le présent rapport propose des définitions pour les *incidents liés à l’IA* et les termes connexes. Ces définitions visent à favoriser une interopérabilité internationale tout en laissant aux juridictions la possibilité de déterminer les dimensions des incidents et *dangers liés à l’IA* qu’elles souhaitent prendre en compte.

1 Executive summary

In January 2023, the OECD formalised the *OECD.AI Expert Group on AI Incidents* to advance the development of *i)* a common AI incident reporting framework and *ii)* an AI Incidents Monitor (AIM). This report provides preliminary definitions and terminology related to AI incidents to support the development and advancement of both initiatives.

As AI systems become more widely used, the potential for them to cause harm to people, organisations and the environment also increases. Harm caused by AI systems can range from minor to severe, and can affect different groups of people, different sectors and different aspects of life.

AI systems need to be trustworthy and reliable to avoid negative effects on people, organisations and the environment. To achieve this, AI actors need to use the same terms to talk about the problems and failures of AI systems so that we can learn at an international level and prevent repeats. These events are broadly referred to under the emerging term *AI incidents*.

This report provides definitions for AI incidents and related terms, based on the work of the OECD.AI Expert Group on AI Incidents and the OECD Working Party on AI Governance (AIGO).

An event where the development or use of an AI system results in *actual* harm is termed an *AI incident*, while an event where the development or use of an AI system is *potentially* harmful is termed an *AI hazard*. This paper defines them as follows:

> *An **AI incident** is an event, circumstance or series of events where the development, use or malfunction of one or more AI systems directly or indirectly leads to any of the following harms:*
> *(a) injury or harm to the health of a person or groups of people;*
> *(b) disruption of the management and operation of critical infrastructure;*
> *(c) violations of human rights or a breach of obligations under the applicable law intended to protect fundamental, labour and intellectual property rights;*
> *(d) harm to property, communities or the environment.*
>
> *An **AI hazard** is an event, circumstance or series of events where the development, use or malfunction of one or more AI systems could plausibly lead to an AI incident, i.e., any of the following harms:*
> *(a) injury or harm to the health of a person or groups of people;*
> *(b) disruption of the management and operation of critical infrastructure;*
> *(c) violations to human rights or a breach of obligations under the applicable law intended to protect fundamental, labour and intellectual property rights;*
> *(d) harm to property, communities or the environment.*

This report also includes proposed definitions for associated terminology, including what constitutes *AI hazards*, *serious AI hazards*, *serious AI incidents* and *AI disasters*, without being overly prescriptive. These definitions are designed to facilitate international interoperability, provide the flexibility necessary to encompass *actual* and/or *potential* harms and allow each jurisdiction to determine the range of AI incidents and hazards they wish to address.

2 Background and purpose

At the second meeting of the OECD Working Party on Artificial Intelligence Governance (AIGO) in November 2022, the OECD Secretariat presented a concept note on developing a common AI incident reporting framework, including the goals of a successful common framework at the OECD level and beyond. The note summarised consultations and preliminary work to date on the topic and presented a preliminary working definition of an "AI incident" considered by the Secretariat and the OECD Expert Group on AI Classification & Risk.

This document contains a draft definition of an AI incident and related terminology. It builds on that concept note and the documents "Initial stocktaking of AI incident definitions and related terminology" [DSTI/CDEP/AIGO/RD(2023)1], "Stocktaking for the development of an AI incident definition" [DSTI/CDEP/AIGO(2022)11/REV1 and DSTI/CDEP/AIGO(2022)11/REV2], presented to AIGO in November 2022, April 2023 and July 2023, respectively. This document was discussed at the 5th and 6th AIGO meetings in November 2023 and February 2024, and incorporates feedback from national delegations and members of the OECD.AI expert group on AI incidents.

This document aims to inform the development of a common AI incidents reporting framework and the AI Incidents Monitor (AIM).

3 *Actual vs. potential* harm as the starting point

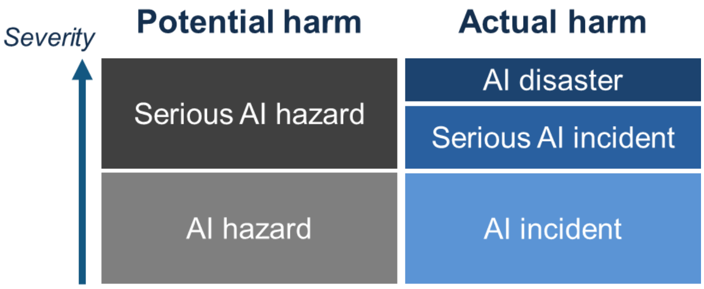
The concept of *harm* is central to the technical standards and regulations that define incidents and hazards. Different frameworks consider different dimensions of harm depending on the specific context, regulatory environment, goals and areas of impact (Annex A). A common thread among most of these frameworks is that incident definitions they use often focus on *potential* harm, *actual* harm or both (OECD, 2023[1]).

Potential harm is often expressed as the risk or likelihood that harm or damage will occur. Risk is a function of both the probability of an event occurring and the severity of the consequences that would result. For example, the risk of an explosion in a chemical plant is greater if the plant is in a densely populated area, and the consequences of an explosion would be severe. It is crucial to identify and address risks and hazards that can arise from the development and use of AI systems for risk management and AI incident reporting frameworks. Potential harm is commonly associated with the concept of *hazard* (OECD, 2023[1]).

Actual harm is often expressed as a risk that materialised into harm. Definitions of actual harm in standards and regulations depend significantly on context. They generally focus on physical injury or damage to health, property or the environment. Some standards and regulations, such as the European Union's General Data Protection Regulation (GDPR), employ the term *damage* to refer to harms (Regulation 2016/679, EU[2]). Actual harm is often associated with the concept of an *incident* (OECD, 2023[1]).

Figure 1 details the proposed OECD classification of AI incidents and hazards based on the severity of harm. This classification is based on a stocktaking exercise of over 30 frameworks, standards and legal instruments relevant to risk and harms (OECD, 2023[1]). It was further informed by discussions with the OECD Working Group on AI Governance (AIGO) and numerous workshops of the OECD.AI expert group on AI incidents.

Figure 1. Proposed classification for AI incidents and hazards based on severity of harm



Source: Adapted from OECD (2023[1]) and discussions with the OECD.AI expert group on AI incidents.

The following sections provide preliminary definitions for the AI incidents and hazards terminology for potential and actual harm.

4 Actual harm: AI incidents, serious AI incidents and AI disasters

Draft definition of an AI incident

> *An **AI incident** is an event, circumstance or series of events where the development, use or malfunction of one or more AI systems directly or indirectly leads to any of the following harms:*
>
> *(a) injury or harm to the health of a person or groups of people;*
>
> *(b) disruption of the management and operation of critical infrastructure;*
>
> *(c) violations of human rights or a breach of obligations under the applicable law intended to protect fundamental, labour and intellectual property rights;*
>
> *(d) harm to property, communities or the environment.*

This working definition of an AI incident is based on the *Concept note on developing a common framework for AI incident reporting and an AI incidents monitor* (OECD, 2022[3]), the definition of a serious AI incident being proposed in the context of the EU AI Act, and numerous discussions with the OECD Working Party on Artificial Intelligence Governance (AIGO) and the OECD.AI expert group on AI incidents.

Key clarifications:

- AI incidents could result in harm to individuals, groups, organisations, communities, society, and the environment.
- For the avoidance of doubt, an event or series of events could include, inter alia, AI system malfunctions and incidents arising from the interaction between two or more AI systems, including agentic AI systems.
- In some cases, the development of an AI system may cause harms even before the system is broadly deployed. For example, training an AI model with proprietary information could infringe copyright laws.
- *Use* includes harms arising from uses of the AI system outside of its intended purposes and intentional or unintentional misuse.
- *Groups of people* includes the concept of *community*, which refers to people living in the same place, area, etc. or having a particular characteristic or activity in common.
- Psychological harms and harms to mental health are included under the broader concept of *health* in (a).

- In some jurisdictions, critical infrastructure is related to *critical functions*. These commonly include both physical and non-physical infrastructure and functions, such as the financial and electoral systems.
- Reputational harm to individuals and intangible harms such as hate speech and mis- and disinformation are included under (c) in relation to a breach of *fundamental rights*.
- Harms to democratic processes may relate to (b) for countries where conducting elections is considered a critical infrastructure or function and under the concept of *harm to communities* in (d) for other countries.
- Violations of intellectual property rights and copyright may fall under both (c) and (d). For example, violations of authors' and *persona*l rights fit into (c), while violations of industrial property rights are included under the concept of *harm to property* in (d).
- Reputational harms to organisations as well as financial harms are included in (d) under the concept of *harm to property*.

Draft definition of a serious AI incident

> *A **serious AI incident** is an event, circumstance or series of events where the development, use or malfunction of one or more AI systems directly or indirectly leads to any of the following harms:*
>
> *(a) the death of a person or serious harm to the health of a person or groups of people;*
>
> *(b) a serious and irreversible disruption of the management and operation of critical infrastructure;*
>
> *(c) a serious violation of human rights or a serious breach of obligations under the applicable law intended to protect fundamental, labour and intellectual property rights;*
>
> *(d) serious harm to property, communities or the environment.*

This working definition of a *serious AI incident* aligns with the definition proposed in the context of the EU AI Act (Annex B). Serious AI incidents are a subset of AI incidents and are intended to include substantial, material incidents.

Key clarifications:

- Assessing the *seriousness* of an AI incident is highly context-dependent; each jurisdiction may define it differently.
- In some cases, the accumulation of smaller AI incidents could lead to a serious AI incident. The wording "event, circumstance or series of events" aims to account for this possibility.
- In some cases, harms inflicted by serious AI incidents can be prolonged over time.

Draft definition of an AI disaster

> *An **AI disaster** is a serious AI incident that **disrupts** the functioning of a community or a society and that may test or exceed its capacity to cope, using its own resources. The effect of an AI disaster can be immediate and localised, or widespread and lasting for a long period of time.*

This working definition of an AI disaster is based on the definitions of disaster by the United Nations Office for Disaster Risk Reduction and the International Federation of Red Cross and Red Crescent Societies (UNDRR, 2023[4]; IFRC, 2023[5]). AI disasters are a subset of serious AI incidents.

5 Potential harm: AI hazards and serious AI hazards

Draft definition of an AI hazard

> *An **AI hazard** is an event, circumstance or series of events where the development, use or malfunction of one or more AI systems could plausibly lead to an AI incident, i.e., any of the following harms:*
>
> *(a) injury or harm to the health of a person or groups of people;*
>
> *(b) disruption of the management and operation of critical infrastructure;*
>
> *(c) violations to human rights or a breach of obligations under the applicable law intended to protect fundamental, labour and intellectual property rights;*
>
> *(d) harm to property, communities or the environment.*

This working definition of an AI hazard is based on the *Concept note on developing a common framework for AI incident reporting and AI incidents monitor* (OECD, 2022[3]), the definition proposed in the context of the EU AI Act, and numerous discussions with the OECD Working Party on Artificial Intelligence Governance (AIGO) and the OECD.AI expert group on AI incidents.

Key clarifications:

- *Near misses* are events that could have led to an AI incident and are therefore included under this definition of AI hazards.
- AI-related risks are included under AI hazards, to the extent that such risks could lead to AI incidents.
- AI hazards could result in harm to individuals, groups, organisations, communities, society, and the environment.
- For the sake of clarity, an event, circumstance or series of events could include, inter alia, AI system malfunctions as well as interactions between two or more AI systems, including agentic AI systems.
- In the context of AI, hazards are not simply AI models in general but also elements of the design, training, and operating context of the AI system. Hazards related to AI systems may be present at any stage of the AI system lifecycle, as modelled in the OECD Framework for the Classification of AI Systems (OECD, 2022[6]).

Draft definition of a serious AI hazard

> *A **serious AI hazard** is an event, circumstance or series of events where the development, use or malfunction of one or more AI systems could plausibly lead to a serious AI incident or AI disaster, i.e., any of the following harms:*
>
> *(a) the death of a person or serious harm to the health of a person or groups of people;*
>
> *(b) a serious and irreversible disruption of the management and operation of critical infrastructure;*
>
> *(c) a serious violation of human rights or a serious breach of obligations under the applicable law intended to protect fundamental, labour and intellectual property rights;*
>
> *(d) serious harm to property, communities or the environment;*
>
> *(e) the disruption of the functioning of a community or a society and which may test or exceed its capacity to cope using its own resources.*

Serious AI hazards are a subset of AI hazards.

Key clarifications:

- Assessing the *seriousness* of an AI hazard is highly context dependent – each jurisdiction may define it differently.
- In some cases, the accumulation of smaller AI hazards could lead to a serious AI hazard. The wording "event, circumstance or series of events" aims to account for this possibility.

6 Conclusion and next steps

The OECD.AI expert group on AI incidents aims to develop a common AI incident reporting framework to enable the appropriate assessment of harms and risks in the AI context. The framework will help identify the key *types* of harm, such as physical, environmental, economic and reputational harm, harm to public interest and harm to fundamental rights. It will also address further *dimensions* of harm, such as level of severity, scope, geographic scale, tangibility, quantifiability, materialisation, reversibility, recurrence, impact and timeframe (Annex A).

A further step would be to establish clear taxonomies to categorise incidents for each dimension of harm. Assessing the “seriousness” of an AI incident, harm, damage, or disruption (*e.g.,* to determine whether an event is classified as an *incident* or a *serious incident*) is context-dependent and is also left for further discussion.

The definitions included in this paper are proposed to serve as the basis for the development and advancement of the common AI reporting framework and the OECD AI Incidents Monitor (AIM).

Annex A. Dimensions of harm

Defining harm and assessing its types, severity levels and other relevant dimensions (*e.g.,* scope, geographic scale, quantifiability, etc.) is key to identifying the incidents that lead or might lead to that harm, and to elaborate an effective framework to address them. Harm definitions and taxonomies are often context-specific. For the purposes of this paper, harm is an umbrella concept that encompasses both the risks of harm (*i.e.*, potential harm) and their materialisation (*i.e.*, actual harm).

Harms can be of different types (*e.g.,* physical, psychological, economic, etc.) and have different levels of severity (*e.g.,* from inconsequential hazards to damage to property, to harm to health, to impact to critical infrastructure, to causing human deaths, etc.). Certain aspects of harm may be quantifiable, such as financial loss or number of impacted individuals. Others may be harder to quantify, such as reputational harm. Harm can be tangible, such as physical injury to a person, or damage to property or the environment. Some harms, such as psychological harms, may not be as tangible or readily quantifiable. Other dimensions of harm include its possible recurrence and reversibility.

The scope and geographical scale of harm are also important. For example, EUROPOL highlights the possibility of large language models being used to “facilitate the perpetration of disinformation, hate speech and terrorist content online”, in addition to providing false objectivity to the messages, and at significantly expanded scale (EUROPOL, 2023[7]).

This indicates the possibility of harmful effects from the development and use of AI, impacting not only individuals but also specific groups or society as a whole. AI has the capability to exacerbate existing problems, as seen in the use of algorithms on social media and may introduce new issues. For example, AI may amplify negative mental health effects, erosion of ethical and cultural values, societal division, and manipulation of electoral preferences.

Table 1 illustrates the dimensions of harm identified from an analysis of over 30 frameworks and legislative instruments (OECD, 2023[1]). These dimensions provide a baseline for further development and discussion on the specificities that an AI incident reporting framework should contain.

Table 1. Illustrative dimensions of harm

Dimensions of harm	Potential criteria for classification
Type	Physical, psychological, reputational, economic/financial (including harm to property), environmental, public interest (*e.g.,* protection of critical infrastructure and democratic institutions), human rights and fundamental rights
Level of severity	Hazard, incident, serious incident, accident, catastrophe; low, medium, high; minor, major, critical; numeric or alphabetical scale
Scope (type of harmed entity)	Individual, group, organisation, institution, society, environment, property
Geographic scale	Single entity, local, national, regional, global
Tangibility	Tangible, intangible
Quantifiability	Quantifiable, unquantifiable
Materialisation	Potential harm (not materialised *e.g.,* hazard), actual harm (materialised *e.g.,* serious incident)
Reversibility	Harm is reversible/irreversible
Recurrence	One-off harms, cumulative effects
Impact	Direct (to individuals), indirect (*e.g.,* to a group, society, environment or public interest; externalities)
Timeframe	Short, medium, long term; within a certain period

Source: OECD (2023[1]).

The following section provides an initial exploration of the current landscape for types of harm.

Types of harm

One of the most relevant dimensions of harm when defining an incident is its type. Specific types of harm are included in sectoral and horizontal technical standards and regulations, depending on the goals, context and industry (OECD, 2023[1]). An AI incident may result in one or multiple of the following types of harm:

- ***Physical harm***: In standards related to product safety or functional safety, physical harm can be categorised according to the type or severity of the injury. For example, the IEC 60950-1 standard for information technology equipment defines physical injury categories as "slight," "moderate," and "severe" (International Electrotechnical Commission, 2010[8]).
- ***Environmental harm***: Some standards categorise harm based on the type of environmental damage caused, such as soil contamination, air pollution, or water pollution. For example, the ISO 14001 standard for environmental management systems includes categories for "minor environmental impact" and "major environmental impact" (International Organization for Standardization, 2015[9]).
- ***Economic or financial harm, including harm to property***: In standards related to financial or economic risk, harm can be categorised based on the magnitude of financial loss or damage. For example, the Basel Framework provides standardised approaches to risk management in the banking sector, addressing risks to credit, market, and operation. (Basel Committee on Banking Supervision, 2017[10]).
- ***Reputational harm***: In standards related to business or organisational risk, harm can be categorised based on the potential impact to an organisation’s reputation or public trust in that an organisation. For example, the ISO 26000 standard for social responsibility includes categories for "minor," "moderate," and "major" negative impacts on reputation (International Organization for Standardization, 2010[11]). Individuals may also be affected by reputational harm (European Union, 2007[12]).
- ***Harm to public interest:*** The International Society of Automation provides the ISA/IEC 62443 Series of Standards, which account for cybersecurity risks that may cause harm to critical infrastructure. It defines levels of security, reliability and integrity (International Society of Automation, 2009[13]). Harm to public interest includes harms to critical infrastructure and functions such as the political system and the rule of law. It also includes harms to the social fabric of a communities.
- ***Harm to human rights and to fundamental rights:*** These rights are established in domestic and international law (United Nations, 1948[14]; European Union, 2007[12]). The EU General Data Protection Regulation (GDPR) is a well-known example of a regulation requiring that certain companies carry out impact assessments to identify and manage risks that may cause harm to privacy rights and other fundamental rights and freedoms of natural persons (Regulation 2016/679, EU[2]).
- ***Psychological harm***: Increasing inclusion of psychological harm and harm to mental health in standards and product safety legislations reflects a growing recognition of the need to consider the full range of potential impacts of products, services, and business operations on individuals and communities (Children Act 1989, UK[15]; The Children Order 1995, Northern Ireland[16]; Scottish Government, 2021[17]; European Parliament, 2024[18]). The concept of psychological harm can be more difficult to assess and quantify than physical harm.

Annex B. Serious AI incident definition in the proposed EU AI Act

The definition of a serious AI incident included in the latest proposal of the EU AI Act is:

> *'Serious incident' means any incident or malfunctioning of an AI system that directly or indirectly leads to any of the following:*
>
> *(a) the death of a person or serious harm to a person's health;*
>
> *(b) a serious and irreversible disruption of the management and operation of critical infrastructure.*
>
> *(c) the infringement of obligations under Union law intended to protect fundamental rights;*
>
> *(d) serious harm to property or the environment;*

Note: This text was taken from the version of the proposed EU AI Act that the European Parliament plenary adopted on 13 March 2024, article 3, point (49) (European Parliament, 2024[18]).

References

Basel Committee on Banking Supervision (2017), *Bank for International Settlements*, https://www.bis.org/bcbs/basel3.htm. [10]

Children Act 1989 (UK), *Children Act 1989 (UK), c. 41*, https://www.legislation.gov.uk/ukpga/1989/41/section/31A. [15]

European Parliament (2024), *Artificial Intelligence Act, European Parliament legislative resolution of 13 March 2024*, https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.pdf. [18]

European Union (2007), *Treaty of Lisbon: Amending the Treaty on European Union and the Treaty Establishing the European Community, 13 December 2007, 2007/C 306/01*, https://www.refworld.org/docid/476258d32.html. [12]

EUROPOL (2023), *ChatGPT - the impact of Large Language Models on Law Enforcement*, Europol Public Information. [7]

IFRC (2023), *What is a disaster?*, https://www.ifrc.org/our-work/disasters-climate-and-crises/what-disaster (accessed on 11 August 2023). [5]

International Electrotechnical Commission (2010), *IEC 61508:2010 CMV - Functional safety of electrical/electronic/programmable electronic safety-related systems*, https://webstore.iec.ch/publication/22273. [8]

International Organization for Standardization (2015), *ISO 14001:2015 Environmental management systems — Requirements with guidance for use*, https://www.iso.org/standard/60857.html. [9]

International Organization for Standardization (2010), *ISO 26000 Guidance on social responsibility*, https://www.iso.org/standard/42546.html. [11]

International Society of Automation (2009), *ISA/IEC 62443 Series of Standards*, https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards. [13]

OECD (2023), “Stocktaking for the development of an AI incident definition”, *OECD Artificial Intelligence Papers*, No. 4, OECD Publishing, Paris, https://doi.org/10.1787/c323ac71-en. [1]

OECD (2022), *DSTI-CDEP-AIGO(2022)11: Concept note on developing a common framework for AI incident reporting and AI incidents monitor*. [3]

OECD (2022), “OECD Framework for the Classification of AI systems”, *OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, https://doi.org/10.1787/cb6d9eca-en. [6]

Regulation 2016/679 (EU), *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Da*, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN#d1e1374-1-1. [2]

Scottish Government (2021), *National Guidance for Child Protection in Scotland*, https://www.gov.scot/binaries/content/documents/govscot/publications/advice-and-guidance/2021/09/national-guidance-child-protection-scotland-2021/documents/national-guidance-child-protection-scotland-2021/national-guidance-child-protection-scotland-2021/g. [17]

The Children Order 1995 (Northern Ireland), *The Children (Northern Ireland) Order 1995 No. 755 (N.I. 2)*, https://www.legislation.gov.uk/nisi/1995/755/article/2/made. [16]

UNDRR (2023), *Disaster: Sendai Framework Terminology On Disaster Risk Reduction*, https://www.undrr.org/terminology/disaster (accessed on 11 August 2023). [4]

United Nations (1948), *Universal Declaration of Human Rights, 10 December 1948, 217 A (III)*, https://www.refworld.org/docid/3ae6b3712c.html. [14]