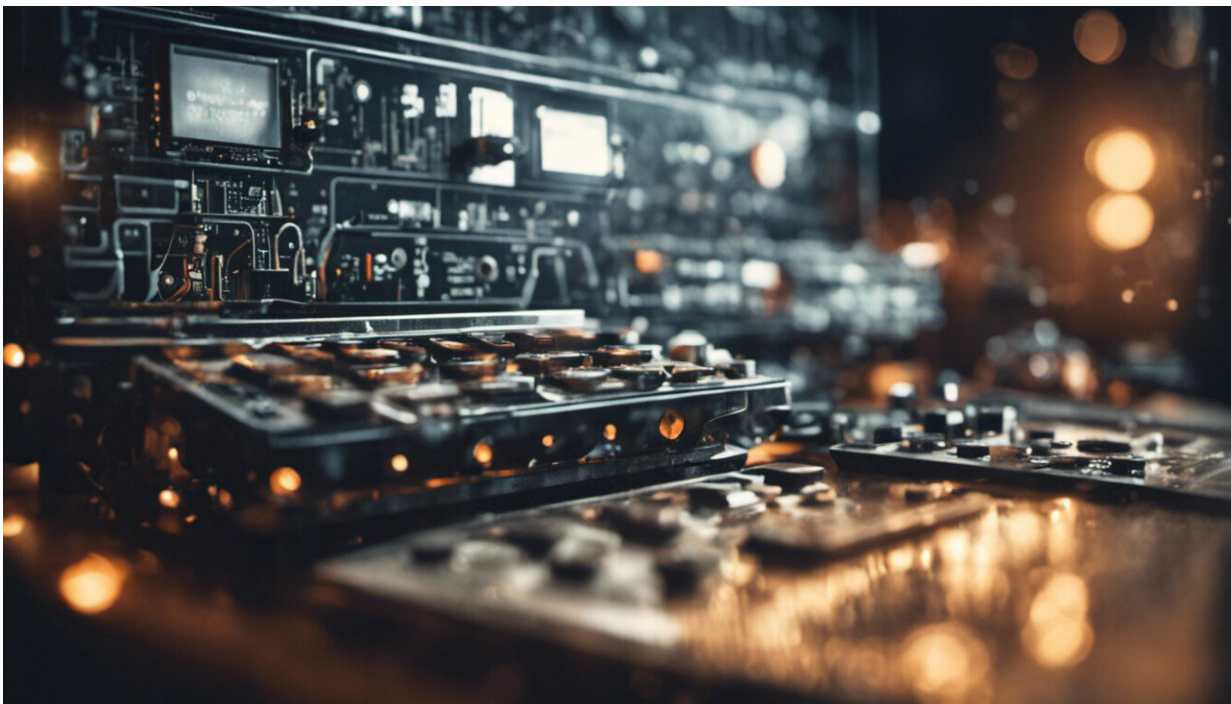


# Opinion: We need to prepare for the public safety hazards posed by artificial intelligence

May 15 2023, by Ali Asgary

---



Credit: AI-generated image ([disclaimer](#))

For the most part, the focus of contemporary emergency management has been on [natural, technological and human-made hazards](#) such as flooding, earthquakes, tornadoes, industrial accidents, extreme weather events and cyber attacks.

However, with the increase in the availability and capabilities of artificial intelligence, we may soon see emerging public safety hazards related to these technologies that we will need to mitigate and prepare for.

Over the past 20 years, my colleagues and I—along with many other researchers—have been [leveraging AI](#) to [develop models](#) and [applications that can](#) identify, assess, predict, monitor and detect hazards to [inform emergency response operations and decision-making](#).

We are now reaching a turning point where AI is becoming a potential source of risk at a scale that should be incorporated into risk and emergency management phases—mitigation or prevention, preparedness, response and recovery.

## **AI and hazard classification**

AI hazards can be classified into two types: intentional and unintentional. Unintentional hazards are those caused by [human errors or technological failures](#).

As the use of AI increases, there will be more adverse events caused by human error in AI models or technological failures in AI based technologies. These events can occur in all kinds of industries including transportation (like drones, trains or self-driving cars), electricity, oil and gas, finance and banking, agriculture, health and mining.

Intentional AI hazards are potential threats that are caused by using AI to harm people and properties. AI can also be used to gain unlawful benefits by compromising security and [safety systems](#).

In my view, this simple intentional and unintentional classification may not be sufficient in case of AI. Here, we need to add a new class of

emerging threats—the possibility of AI overtaking human control and decision-making. This may be triggered intentionally or unintentionally.

Many AI experts have already warned [against such potential threats](#). A recent open letter by researchers, scientists and others involved in the development of AI [called for a moratorium on its further development](#).

## **Public safety risks**

Public safety and emergency management experts use risk matrices to assess and compare risks. Using this method, hazards are qualitatively or quantitatively assessed based on their frequency and consequence, and their impacts are classified as low, medium or high.

Hazards that have low frequency and low consequence or impact are considered low risk and no additional actions are required to manage them. Hazards that have medium consequence and medium frequency are considered medium risk. These risks need to be closely monitored.

Hazards with high frequency or high consequence or high in both consequence and frequency are classified as high risks. These risks need to be reduced by taking additional risk reduction and mitigation measures. Failure to take immediate and proper action may result in severe human and property losses.

Up until now, AI hazards and risks have not been added into the risk assessment matrices much beyond organizational use of AI applications. The time has come when we should quickly start bringing the potential AI risks into local, national and global risk and emergency management.

### **AI risk assessment**

AI technologies are becoming more widely used by institutions,

organizations and companies in different sectors, and hazards associated with the AI are starting to emerge.

In 2018, the accounting firm KPMG developed an "[AI Risk and Controls Matrix](#)." It highlights the risks of using AI by businesses and urges them to recognize these new emerging risks. The report warned that AI technology is advancing very quickly and that risk control measures must be in place before they overwhelm the systems.

Governments have also started [developing some risk assessment guidelines](#) for the use of AI-based technologies and solutions. However, these guidelines are limited to risks such as algorithmic bias and violation of individual rights.

At the government level, the Canadian government issued the "[Directive on Automated Decision-Making](#)" to ensure that federal institutions minimize the risks associated with the AI systems and create appropriate governance mechanisms.

The main objective of the directive is to ensure that when AI systems are deployed, risks to clients, federal institutions and Canadian society are reduced. According to this directive, risk assessments must be conducted by each department to make sure that appropriate safeguards are in place in accordance with the [Policy on Government Security](#).

In 2021, the U.S. Congress tasked the National Institute of Standards and Technology with developing an AI risk management framework for the Department of Defense. The proposed voluntary AI [risk assessment](#) framework recommends banning [the use of AI systems that present unacceptable risks](#).

## Threats and competition

Much of the national level policy focus on AI has been from [national security](#) and global competition perspectives—the national security and economic risks of falling behind in the AI technology.

The U.S. National Security Commission on Artificial Intelligence highlighted [national security risks associated with AI](#). These were not from the public threats of the technology itself, but from losing out in the global competition for AI development in other countries, including China.

In its 2017 [Global Risk Report](#), the World Economic Forum highlighted that AI is only one of emerging technologies that can exacerbate global risk. While assessing the risks posed by the AI, the report concluded that, at that time, super-intelligent AI systems remain a theoretical threat.

However, the latest [Global Risk Report 2023](#) does not even mention the AI and AI associated risks which means that the leaders of the global companies that provide inputs to the global risk report had not viewed the AI as an immediate risk.

## **Faster than policy**

AI development is progressing much faster than government and corporate policies in understanding, foreseeing and managing the risks. The current global conditions, combined with market competition for AI technologies, make it difficult to think of an opportunity for governments to pause and develop risk governance mechanisms.

While we should collectively and proactively try for such governance mechanisms, we all need to brace for major catastrophic AI's impacts on our systems and societies.

This article is republished from [The Conversation](#) under a Creative

Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Opinion: We need to prepare for the public safety hazards posed by artificial intelligence (2023, May 15) retrieved 23 April 2024 from <https://techxplore.com/news/2023-05-opinion-safety-hazards-posed-artificial.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.