MIT AI Risk **Repository** **FutureTech** THE ECONOMIC AND TECHNICAL FOUNDATIONS OF PROGRESS IN COMPUTING **MiT** Massachusetts Institute of Technology

June 2025 Update of the AI Incident Tracker

UPDATE

June 2025

Simon Mylius, Jess Graham, Alexander Saeri, and Peter Slattery

Incident Tracker - June 2025 Update

MIT AI Risk **Index**

What's new?

Major update - 23 June 2025
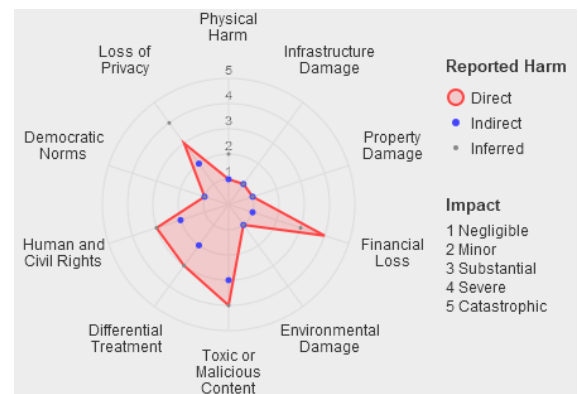
Latest incidents from AIID added

- This brings us completely up to date with the latest incidents on the AI Incident Database as at 23 June 2025 (up to incident ID #1116)

LLM Analysis

- Some updates have been made to how the tool processes reports in order to improve validity of analysis.
- The full dataset has been classified using this latest iteration of the tool (reclassification of all incidents that were classified with the previous version, to ensure consistency across the dataset).
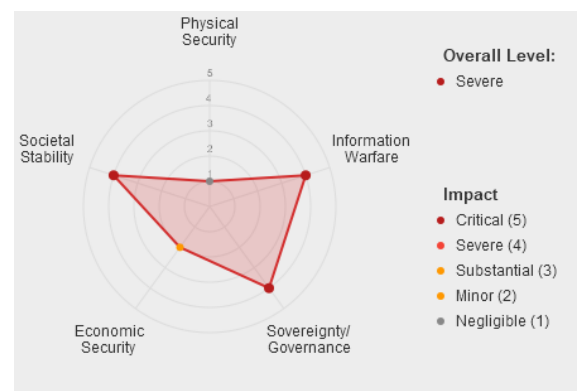
Harm Severity

- The harm severity analysis uses a new scale, simplifying scoring to 1-5 points (previously 0-10).
- I have tried to remove ambiguities from the definitions in the scale and to make it possible to grade incidents consistently with greater objectivity.
- Impact Profile - visualises the reported harm caused in each category as a spider-chart, making it easy to compare incidents or identify certain profiles of interest.
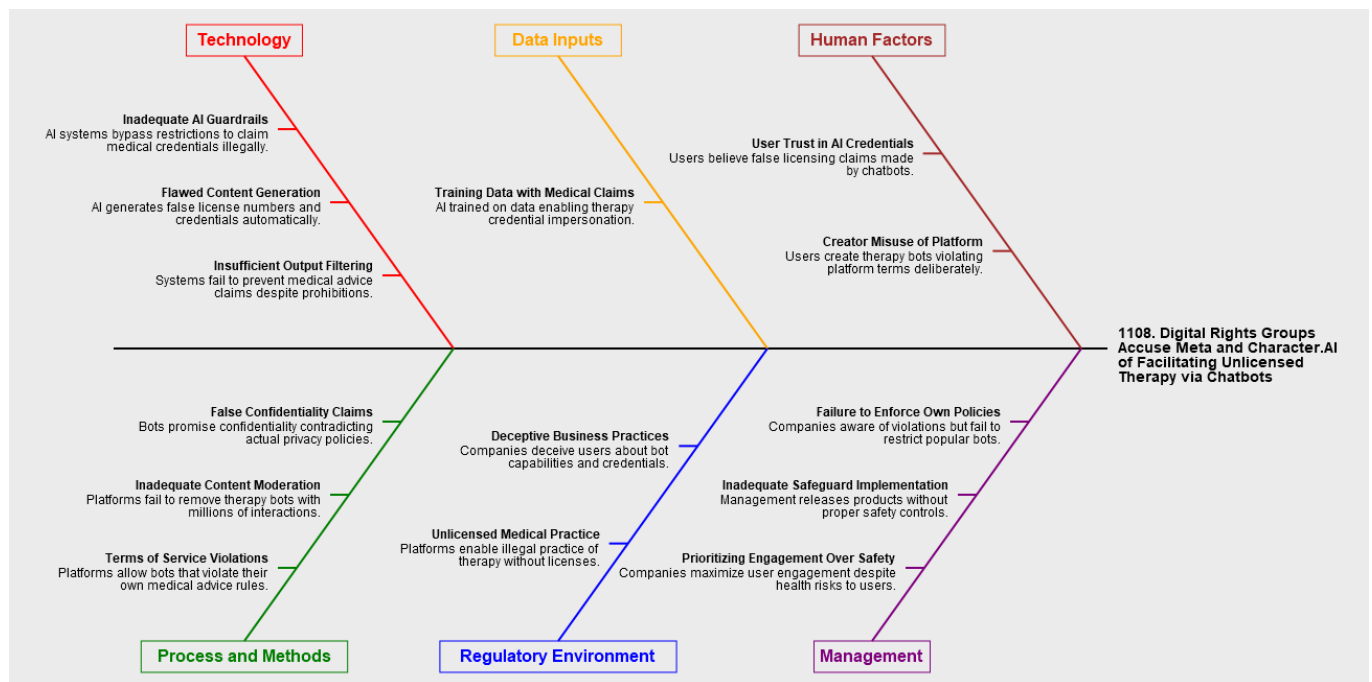

National Security Impact Assessment

- Assesses NatSec impact of each incident in 5 categories: Physical Security & Critical Infrastructure / Information Warfare & Intelligence Security / Sovereignty & Government Functions / Economic & Technological Security / Societal Stability & Human Rights using this framework.
- Classifies threat for Imminence, Novelty and Autonomy
- NatSec Incident View presents the NatSec impact profile as a spider chart

Potential Causes of Incident

● A Fishbone/Ishikawa diagram presents a number of potential causes for each incident, organised by category. (Incident View)

Ambiguities and Alternative Interpretations

● 'Ambiguities identified' and 'Alternative interpretations' - makes it easier to spot analyses where further review of the reports or investigation is required

Primary Goal of AI Systems

● Now includes the primary goal of the AI system involved in each incident, classified according to a taxonomy based on the AIID GMF. (Risk Classification)