Dual-Use Research Norms (DURC for AI)

dual-use-research-taxonomy · safety · concept

Source: https://policywindow.org/wiki/dual-use-research-taxonomy

Generated 2026-11-19T01:06:38 UTC

Summary

A normative framework — adapted from biosecurity's Dual-Use Research of Concern (DURC) policies — for governing AI research and publication decisions when research outputs have both beneficial and harmful applications.

At a glance

Used by

8 instrument(s)

Related concepts

alignment, capability-elicitation, red-team-evaluation, asl-3

Primary source

Solaiman, I., et al. (2019), 'Release Strategies and the Social Impacts of Language Models' — the canonical articulation of structured-access norms for foundation models.

Details

Dual-use research norms in AI explicitly draw on the biosecurity precedent: the 1975 Asilomar conference on recombinant DNA, the 2004 US National Science Advisory Board for Biosecurity, and the 2014 US gain-of-function moratorium. The AI parallels are publication-control debates around GPT-2 (OpenAI's staged release, 2019), the deepfake-generation research community (FaceSwap-era, 2017-2020), CBRN-uplift research, and offensive cybersecurity capabilities (e.g., AutoAttack research). Field positions cluster: (a) full publication — Brundage et al. 2018 critique of selective release; (b) staged or structured access — Solaiman et al. 2019; (c) capability-thresholded redaction — Anthropic, OpenAI, DeepMind dual-use policies, 2023-2025.

Governance instruments are catching up. US EO 14110 §4.2(a)(ii) explicitly required reporting on dual-use capabilities including CBRN, cyber, and autonomous-replication. EU AI Act Art. 5 prohibits certain dual-use applications (manipulation, social scoring) but does not regulate research-stage decisions. NIST AI RMF Map 1.1 includes 'risk of misuse' assessment but does not prescribe publication norms. The G7 Hiroshima Code §3 endorses 'responsible information sharing' without operationalising it.

For AI safety researchers, dual-use research norms are the closest analogue to peer-review-style governance of which findings should be public — a research-community-internal governance layer that operates upstream of regulator-mandated controls.

How to cite this article

APA

Policy Window. (n.d.). Dual-Use Research Norms (DURC for AI) [Wiki article — Concept]. https://policywindow.org/wiki/dual-use-research-taxonomy

CHICAGO

Policy Window. n.d.. "Dual-Use Research Norms (DURC for AI)." Wiki article (Concept). https://policywindow.org/wiki/dual-use-research-taxonomy.

HARVARD

Policy Window (n.d.) 'Dual-Use Research Norms (DURC for AI)', Wiki article — Concept, available at: https://policywindow.org/wiki/dual-use-research-taxonomy.

OSCOLA

Policy Window, 'Dual-Use Research Norms (DURC for AI)' (Wiki article — Concept, n.d.) <https://policywindow.org/wiki/dual-use-research-taxonomy> accessed [date].

BIBTEX

```
@misc{policywindow-dual-use-research-taxonomy,
title  = {Dual-Use Research Norms (DURC for AI)},
author = {Policy Window},
year   = {n.d.},
howpublished = {dual-use-research-taxonomy — safety},
url    = {https://policywindow.org/wiki/dual-use-research-taxonomy},
note   = {Primary source: https://arxiv.org/abs/1908.09203}
}
```