

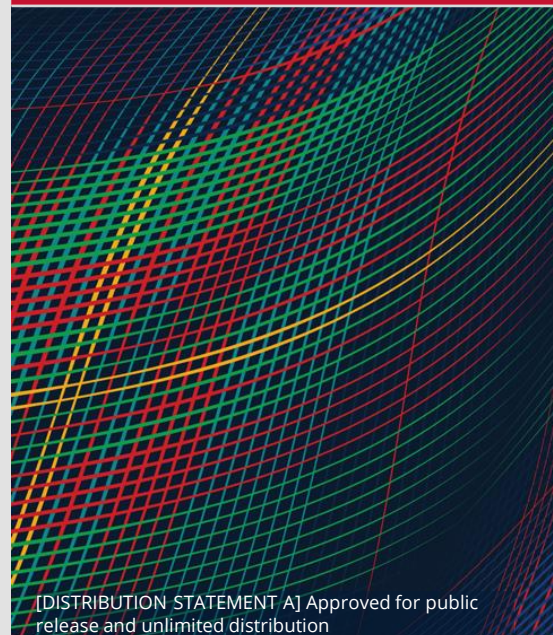


Exploring the AI Incident Documentation Practice

MAY 14, 2026

Violet Turri

**Carnegie
Mellon
University**
Software
Engineering
Institute



Document Markings

Copyright 2026 Carnegie Mellon University.

This material is based upon work funded and supported by the Department of War under Air Force Contract Nos. FA8702-15-D-0002, and FA870225DB003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

The opinions, findings, conclusions, and/or recommendations contained in this material are those of the author(s) and should not be construed as an official US Government position, policy, or decision, unless designated by other documentation.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

This material was prepared for the exclusive use of NIST Workshop and may not be used for any other purpose without the written consent of permission@sei.cmu.edu.

DM26-0516

Challenges Understanding AI Systems and Incidents



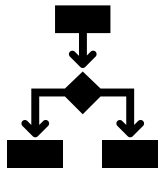
Where: Context-Specific Failures

AI systems and their behaviors are tightly bound to deployment context, data, and model components



When: Factors Throughout the Timeline

AI system lifecycles include many inflection points where things can go wrong, e.g. data updates, model retraining, post-incident responses



Why: Challenges with Explainability

Challenging to determine underlying behaviors and causes from observable characteristics, open issues in explainability

Reporting Mechanisms in High-Stakes Industries

Aviation

- Systematic collection, analysis, investigation, and sharing of incident reports to drive safety improvements
- Reports capture the full timeline from loading to disembarkation
- Options for confident, de-identified reporting

Cyber

- Unique identifiers for publicly known vulnerabilities and standardized vocabularies
- Sharing and analysis across organizations
- Reporting requirements to ensure first-hand, validated data

Opportunity:

Models for centralized, mandatory reporting mechanisms that enable first-hand and/or anonymized accounts with high levels of oversight

Leveraging Databases of AI Systems

Government-run AI System Registries

Provide public, high-level details on system such as:

- System purpose
- Inputs / outputs
- Techniques used
- Responsible parties

Controversial AI Systems Catalogues

Identify systems with known or potential harms, including:

- Location
- Sector
- Ethical concerns

Opportunity

Link systems stored in AI registries to incident reports to help:

- Trace the full lifecycle of an AI system
- Enable proactive risk monitoring

Ideas for Augmenting Existing Practices



Government Oversight:

A mandatory, federally overseen repository would enable first-hand, validated reports that can be investigated



Anonymous Submissions:

A de-identified reporting channel would encourage whistleblowers and engineers to share detailed failure information without fear of repercussion



Proactive Documentation:

Maintaining a database of AI systems *before* incidents occur could help track lifecycles and detect early warning signs