



HISPI Project Cerebellum AI Governance Community of Practice Trusted Artificial Intelligence Model (TAIM) Top 20

July 22, 2026



©2026 HISPI

Disclaimer: The views and opinions expressed in this presentation are mine and **do not** necessarily reflect the official policy or position of any of my former and current employers, clients or affiliates. Assumptions made by me are **not** reflective of the position of any organization or government entity.



Project Cerebellum

AI should cause no harm but enhance the quality of human life

Trusted AI starts with clear guardrails, accountable
decisions, and measured risk.

©2026 HISPI



2

Vision

Provide effective guardrails for AI use cases.

Mission

Harmonize best practices, standards, and frameworks for AI ecosystems.

Trusted AI Model (TAIM)



Risk of AI: Production impact from poor guardrails

A broadly scoped task can become a production incident when access, testing, and rollback are not governed.

What

Agent deleted production data.

Why

Broad token allowed unsafe action.

Lesson

Govern access, testing, and rollback.

Risk control takeaway

AI agents need the same basic safeguards as other high-impact automation: least privilege, separation of environments, change control, logging, rollback, and tested recovery.

©2026 HISPI



4

The Cursor AI coding agent, operating on Anthropic's Claude Opus 4.6, is alleged to have deleted PocketOS's production database and volume-level backups while working on a staging-environment task.

The agent utilized a broadly scoped API token to delete a Railway volume, causing disruption to PocketOS and its car-rental business customers. Railway eventually helped recover data and implemented safety measures post-incident.

AI Incidents: The risk landscape

- **Fraud and deepfakes** Voice clones, fake endorsements, scam ads
- **Bias and access** Unfair outcomes, language barriers, exclusion
- **Privacy and identity** Misuse of names, likeness, personal data
- **Safety and reliability** Bad alerts, unsafe automation, failures
- **Security and supply chain** Third-party systems, agent access, data exposure

Even a narrowly scoped task can become a production incident when access, testing, and rollback are not governed.

©2026 HISPI



5

Trusted AI Model (TAIM) Domains

(NIST AI RMF functions)

TAIM helps organizations introduce AI deliberately and mitigate risk across the AI lifecycle.

GOVERN

Set direction and accountability

MAP

Understand context and risks

MEASURE

Test and evaluate trustworthiness

MANAGE

Respond, improve, and report

The domains create a repeatable pathway: decide what AI should do, understand where it can go wrong, test the trust factors, and manage what happens next.

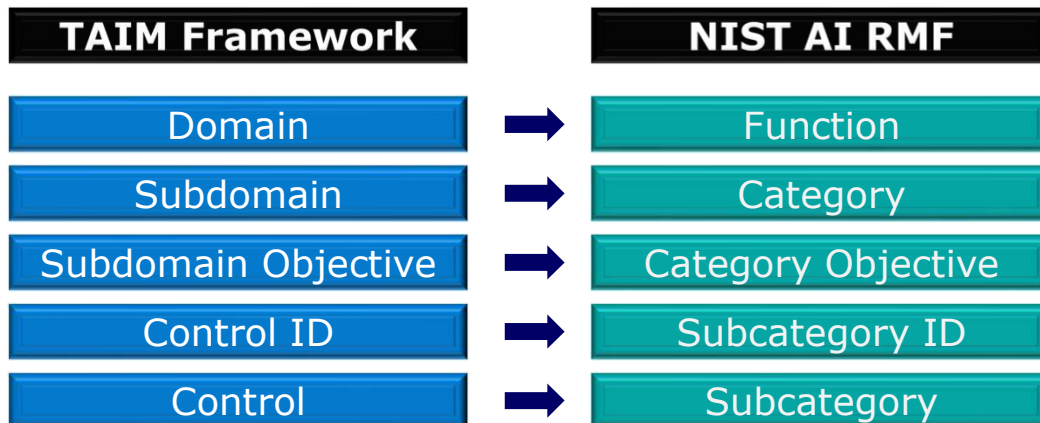


©2026 HISPI

6

Trusted AI Model (TAIM) Taxonomy

(relationship to NIST AI RMF)



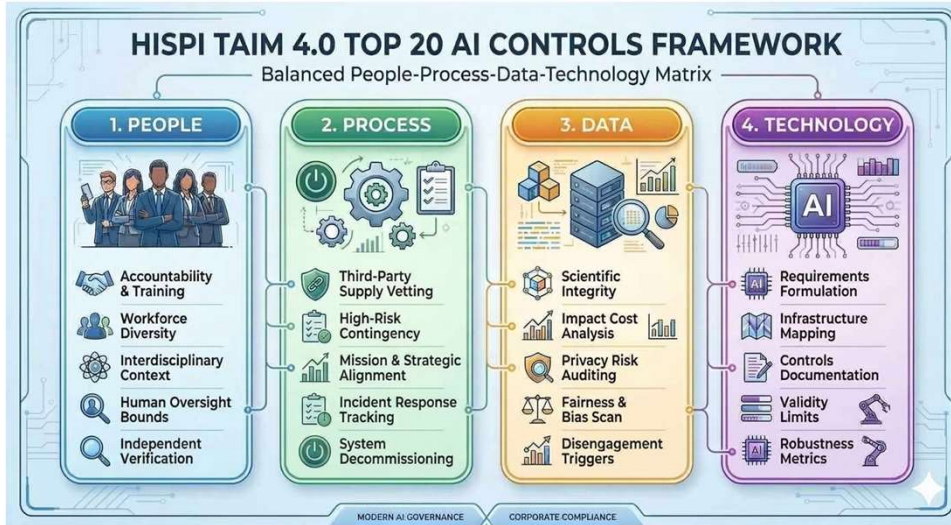
Purpose: give organizations a practical control language while preserving alignment to the NIST Artificial Intelligence Risk Management Framework.

©2026 HISPI



7

TAIM Controls



The TAIM Top 20 is structured into **four core control domains**:

1. People Controls

These controls ensure that the human elements of AI development, use, and oversight are well-managed, trained, and transparent (slide# 6, 10).

Accountability & Training: Ensuring personnel and partners receive documented AI risk management training to perform their lifecycle roles (slide# 12).

Workforce Diversity: Actively prioritizing demographic, disciplinary, and background diversity when mapping, measuring, and managing risks (slide# 13).

Interdisciplinary Context: Documenting the participation and collaboration of diverse AI actors and domain experts to establish system context (slide# 16).

Human Oversight Bounds: Defining, assessing, and documenting explicit human-in-the-loop oversight workflows and technical parameters (slide# 21).

Independent Verification: Involving independent internal experts and external community stakeholders in regular trust and risk assessments (slide# 24).

2. Process Controls

These establish the procedures, methodologies, and operational lifecycles necessary for governing AI models (slide# 6, 10).

Third-Party Supply Vetting: Managing risks associated with external software or data suppliers, including intellectual property infringement (slide# 14).

High-Risk Contingency: Implementing operational contingency workflows to handle immediate failures in high-risk third-party AI dependencies (slide# 15).

Mission & Strategic Alignment: Documenting the organization's core mission, technical use cases, and strategic goals for the AI technology (slide# 17).

Incident Response Tracking: Deploying formal playbooks to track, log, and recover from failures while managing communication pipelines (slide# 30).

System Decommissioning: Establishing formal processes and procedures for phasing out or decommissioning legacy AI systems safely (slide# 11).

3. Data Controls

Because models are only as good as their inputs, these controls address data sourcing, integrity, privacy, and protection (slide# 6, 10).

Scientific Integrity: Documenting upfront Test & Evaluation (TEVV) considerations, data collection suitability, and dataset representativeness (slide# 19).

Impact Cost Analysis: Examining and documenting non-monetary societal or operational costs resulting from expected data or system errors (slide# 20).

Privacy Risk Auditing: Programmatically examining, verifying, and logging data privacy exposure risks over time (slide# 27).

Fairness & Bias Scan: Evaluating programmatic statistical testing results to identify and mitigate distribution or demographic biases (slide# 28).

Disengagement Triggers: Structuring explicit data mechanisms and roles to override, supersede, or fully deactivate misbehaving models (slide# 29).

4. Technology Controls

These focus on the underlying infrastructure, security, and technical safeguards of the deployed AI models (slide# 6, 10).

Requirements Formulation: Eliciting core technical constraints, sociotechnical implications, and privacy rules directly into system specifications (slide# 18).

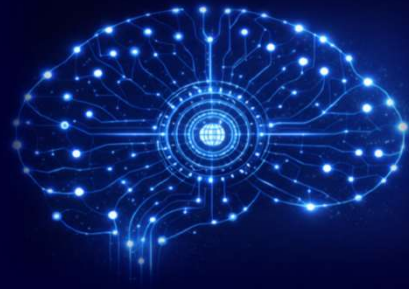
Infrastructure Mapping: Continually tracking component-level vulnerabilities and legal liabilities across proprietary code and external data layers (slide# 22).

Controls Documentation: Cataloging and maintaining specific engineered risk mitigations designed for individual sub-components or APIs (slide# 23).

Validity Limits: Mathematically validating system accuracy while explicitly logging generalizability boundaries past development conditions (slide# 25).

Robustness Metrics: Deploying real-time monitoring and safety metrics to ensure systems fail safely past operational knowledge limits (slide# 26).

TAIM 4.0 Top 20 Controls



©2026 HISPI



9

TAIM Snapshot

A single model view connects domains, objectives, controls, mappings, principles, and supporting evidence.

Domain (Function)	Subdomain (Category ID)	Subdomain Objective (Category Objective)	Control ID (Subcategory ID)	Control (Subcategory)	Control Question	Recommendation	Top 25 Control	Related AI Incidents	PEOPLE	PROCESS	DATA	TECHNOLOGY	ISO/IEC TR 24028:2020 Information Technology - Artificial Intelligence - Overview of Trustworthiness in artificial intelligence	BS ISO/IEC 23894:2022 Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) (British Standard)	ISO/IEC 23894:2022 Information Technology - Artificial Intelligence - Guidance on risk management
GOVERN	GOVERN 2	Managing AI risks	GOVERN 2.2	The organization's personnel and partners receive AI risk management training to enable them to perform their duties and responsibilities consistent with related policies, procedures, and agreements	Do the organization's personnel and partners receive AI risk management training to enable them to perform their duties and responsibilities consistent with related policies, procedures, and agreements?	Ensure that the organization's personnel and partners receive AI risk management training to enable them to perform their duties and responsibilities consistent with related policies, procedures, and agreements	✓	https://ai-incidents.projects.enhelfum.com/tam/govern-2-2	✓	✓	✓	✓	5.4		5.4.1 - ISO 31000:2018
			GOVERN 2.3	Executive leadership of the organization takes responsibility for decisions about risks associated with AI system development and deployment	Does executive leadership of the organization take responsibility for decisions about risks associated with AI system development and deployment?	Ensure that executive leadership of the organization takes responsibility for decisions about risks associated with AI system development and deployment			✓	✓	✓	✓		A.2.2.5	5.5
		Workforce diversity, equity, inclusion, and accessibility processes are consistent with the	GOVERN 3.1	Decision-making related to mapping, measuring, and managing AI risks throughout the lifecycle is informed by a diverse team (e.g., diversity of perspectives, disciplines)	Are decision-making related to mapping, measuring, and managing AI risks throughout the lifecycle informed by a diverse team (e.g., diversity of perspectives, disciplines)?	Ensure that decision-making related to mapping, measuring, and managing AI risks throughout the lifecycle is informed by a diverse team (e.g., diversity of perspectives, disciplines)	✓	https://ai-incidents.projects.enhelfum.com/tam/govern-3-1	✓	✓	✓	✓	7.1	7.2	6.2 - ISO 31000:2018

TRUSTED AI MODEL (TAIM)-V4.0

©2026 HISPI



10

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/govern-1-7	GOVERN 1.7	Processes and procedures are in place for decommissioning and phasing out AI systems safely and in a manner that does not increase risks or decrease the organization’s trustworthiness.	Processes and procedures are in place for decommissioning and phasing out AI systems safely and in a manner that does not increase risks or decrease the organization’s trustworthiness.	BS ISO/IEC 23053:2022 8.1 ISO/IEC DIS 5338:2022 6.3.3.3, 6.4.8.2, 6.4.8.3, 6.4.17.1 ISO/IEC 38507:2022 4.3, 5.3 ISO/IEC 22989:2022 6.1, 6.2.9 ISO/IEC 42001:2023 B.6.2.6 - AI system operation and monitoring Texas: Responsible Artificial Intelligence Governance Act (TRAIGA)		
Applicable Regulation	Risk Statement To ensure a purpose-driven culture focused on risk understanding and management.					
PPDT Governance Framework						
PEOPLE	PROCESS					
DATA	TECHNOLOGY					
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓					✓	



©2026 HISPI

11

Ars Technica Retracted Article After Purportedly AI-Generated Text Was Presented as Direct Quotes From Matplotlib Maintainer

February 13, 2026

Ars Technica reportedly retracted an article due to the misrepresentation of text generated by an AI as direct quotations from a source. The publication acknowledged a breach in standards, stating that the use of AI-generated material did not align with Ars Technica's policy. An apology was issued to readers and the affected party, matplotlib maintainer Scott Shambaugh. It was later disclosed that an author inadvertently used an AI-paraphrased version of the source's text as direct quotes.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all](#)

[TAIM controls](#)

[MAP 1.6](#) — similarity 0.602, rank 1. [TAIM detail and related incidents →](#)

[MAP 4.1](#) — similarity 0.594, rank 2. [TAIM detail and related incidents →](#)

[GOVERN 1.7](#) — similarity 0.593, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/govern-2-2	GOVERN 2.2	Accountability structures are in place so that the appropriate teams and individuals are empowered, responsible, and trained for mapping, measuring, and managing AI risks.	The organization’s personnel and partners receive AI risk management training to enable them to perform their duties and responsibilities consistent with related policies, procedures, and agreements.	ISO/IEC TR 24028:2020 5.4 ISO/IEC 31000:2018 5.4.1 ISO/IEC DIS 5338:2022 6.4.12.1 ISO/IEC 38507:2022 4.1, 4.3, 5.1, 6.7.2, 6.7.5, A.1.3 ISO/IEC 22989:2022 6.2.2 NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 7.2 – Competence Texas: Responsible Artificial Intelligence Governance Act (TRAIGA)		
Applicable Regulation	Risk Statement					
HIPAA PCI-DSS SOC2 EU AI ACT EU GDPR WHITE HOUSE AI EO	To ensure a purpose-driven culture focused on risk understanding and management.					
PPDT Governance Framework						
PEOPLE PROCESS DATA TECHNOLOGY						
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓			✓		

©2026 HISPI



12

Purported Deepfake Scam Ad Reportedly Used Lara Lewington and Martin Lewis to Promote Quantum AI Scheme

March 9, 2026

Scammers allegedly used a fraudulent deepfake video featuring Lara Lewington, coupled with the endorsement of Martin Lewis, to peddle a suspected 'Quantum AI' investment scheme. Reports suggest that this deceptive video was intended to lure viewers into participating in an unverified financial venture.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MAP 1.6](#) — similarity 0.557, rank 1. [TAIM detail and related incidents →](#)

[GOVERN 2.2](#) — similarity 0.553, rank 2. [TAIM detail and related incidents →](#)

[GOVERN 6.2](#) — similarity 0.551, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/govern-3-1	GOVERN 3.1	Workforce diversity, equity, inclusion, and accessibility processes are prioritized in the mapping, measuring, and managing of AI risks throughout the lifecycle.	Decision-making related to mapping, measuring, and managing AI risks throughout the lifecycle is informed by a diverse team (e.g., diversity of demographics, disciplines, experience, expertise, and backgrounds).	ISO/IEC TR 24028:2020 7.1 BS ISO/IEC 23053:2022 7.2 ISO/IEC 31000:2018 6.2 ISO/IEC DIS 5338:2022 6.3.3 ISO/IEC 38507:2022 6.1, 6.3, 6.5, 6.7.1, 6.7.2, 6.7.3, Annex A ISO/IEC 22989:2022 6.2.2 ISO/IEC 42001:2023 B.4.6 - Human resources B.5.4 - Assessing AI system impact on individuals and groups of individuals Texas: Responsible Artificial Intelligence Governance Act (TRAIGA)		
Applicable Regulation	Risk Statement To ensure a purpose-driven culture focused on risk understanding and management.					
PPDT Governance Framework						
PEOPLE PROCESS DATA TECHNOLOGY						
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓	✓	✓			✓

©2026 HISPI

13



13

Sora Video Generator Has Reportedly Been Creating Biased Human Representations Across Race, Gender, and Disability

March 23, 2025

An investigation by WIRED revealed bias in OpenAI's video generation model, Sora. In a test using 250 prompts, Sora was more likely to portray CEOs and professors as men, flight attendants and childcare workers as women, and showed stereotypical depictions of disabled individuals and people with larger bodies.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all](#)

[TAIM controls](#)

[GOVERN 3.1](#) — similarity 0.645, rank 1. [TAIM detail and related incidents →](#)

[MAP 1.6](#) — similarity 0.636, rank 2. [TAIM detail and related incidents →](#)

[MAP 4.1](#) — similarity 0.623, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/govern-6-1	GOVERN 6.1	Policies and procedures are in place to address AI risks and benefits arising from third-party software and data and other supply chain issues.	Policies and procedures are in place that address AI risks associated with third-party entities, including risks of infringement of a third-party's intellectual property or other rights.	ISO/IEC TR 24028:2020 5.5 ISO/IEC 31000:2018 5.4.1 ISO/IEC DIS 5338:2022 6.4.3.3 ISO/IEC 38507:2022 5.5, 6.4, A.3 ISO/IEC 22989:2022 8.6.1 NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 B.10.2 - Allocating responsibilities B.10.3 – Suppliers Texas: Responsible Artificial Intelligence Governance Act (TRAIGA)		
Applicable Regulation	Risk Statement					
HIPAA SOC2 EU GDPR PCI-DSS EU AI ACT WHITE HOUSE AI EO	To ensure a purpose-driven culture focused on risk understanding and management.					
PPDT Governance Framework						
PEOPLE DATA PROCESS TECHNOLOGY						
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓		✓	✓	✓	

©2026 HISPI

14

Grammarly's AI Expert Review Allegedly Used Journalists' and Authors' Names Without Consent

March 11, 2026

Grammarly's Expert Review feature, utilizing a large language model, is accused of generating editing suggestions under the names of journalists, authors, and academics without their approval. The federal class action led by Julia Angwin asserts this practice misappropriates identities for commercial purposes, attributing advice never given by the named individuals.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all](#)

[TAIM controls](#)

[GOVERN 1.1](#) — similarity 0.643, rank 1. [TAIM detail and related incidents →](#)

[MAP 4.1](#) — similarity 0.642, rank 2. [TAIM detail and related incidents →](#)

[GOVERN 6.1](#) — similarity 0.637, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk									
https://ai-incidents.projectcerebellum.com/taim/govern-6-2	GOVERN 6.2	Policies and procedures are in place to address AI risks and benefits arising from third-party software and data and other supply chain issues.	Contingency processes are in place to handle failures or incidents in third-party data or AI systems deemed to be high-risk.	ISO/IEC TR 24028:2020 7.1 BS ISO/IEC 23053:2022 8.3 ISO/IEC 31000:2018 5.4.1 ISO/IEC DIS 5338:2022 6.4.15.3 ISO/IEC 38507:2022 6.7.5 ISO/IEC 22989:2022 5.15.4 NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 B.10.2 - Allocating responsibilities B.10.3 – Suppliers Texas: Responsible Artificial Intelligence Governance Act (TRAIGA)									
Applicable Regulation <table><tr><td>HIPAA</td><td>PCI-DSS</td></tr><tr><td>SOC2</td><td>EU AI ACT</td></tr><tr><td>EU GDPR</td><td>WHITE HOUSE AI EO</td></tr></table> PPDT Governance Framework <table><tr><td>PEOPLE</td><td>PROCESS</td></tr><tr><td>DATA</td><td>TECHNOLOGY</td></tr></table>	HIPAA	PCI-DSS	SOC2	EU AI ACT	EU GDPR	WHITE HOUSE AI EO	PEOPLE	PROCESS	DATA	TECHNOLOGY	Risk Statement To ensure a purpose-driven culture focused on risk understanding and management.		
HIPAA	PCI-DSS												
SOC2	EU AI ACT												
EU GDPR	WHITE HOUSE AI EO												
PEOPLE	PROCESS												
DATA	TECHNOLOGY												
Core Principles													
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)							
✓	✓		✓	✓	✓								

©2026 HISPI

15

Purported Deepfake Impersonation of Elon Musk Used to Promote Fraudulent '17-Hour' Diabetes Treatment Claims

December 27, 2025

A deepfake video falsely attributed to Elon Musk promoted an unverified '17-hour' diabetes cure. The misleading content appeared as part of a scam ecosystem, highlighting the need for trustworthy and safe AI practices. Rapper Boosie Badazz unwittingly amplified the video before its authenticity was questioned.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MEASURE 2.10](#) — similarity 0.569, rank 1. [TAIM detail and related incidents →](#)

[MAP 1.6](#) — similarity 0.563, rank 2. [TAIM detail and related incidents →](#)

[GOVERN 6.2](#) — similarity 0.559, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk														
https://ai-incidents.projectcerebellum.com/taim/map-1-2	MAP 1.2 Risk Statement To ensure sufficient contextual knowledge about AI system impacts to inform an initial go/no-go decision about whether to design, develop, or deploy an AI system.	Context is established and understood.	Interdisciplinary AI actors, competencies, skills, and capacities for establishing context reflect demographic diversity and broad domain and user experience expertise, and their participation is documented. Opportunities for interdisciplinary collaboration are prioritized.	ISO/IEC TR 24028:2020 9.8 ISO/IEC DIS 5338:2022 5.1, 6.2.3.3, 6.2.4.1, 6.2.4.3, 6.2.6.1, 6.2.6.3, 6.4.9.4 ISO/IEC 38507:2022 A.1.3 ISO/IEC 22989:2022 5.1, 6.2.2 NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 B.4.6 - Human resources 7.2 – Competence Utah: AI Policy Act (2024)														
Core Principles																		
Applicable Regulation HIPAA PCI-DSS SOC2 EU AI ACT EU GDPR WHITE HOUSE AI EO																		
PPDT Governance Framework PEOPLE PROCESS DATA																		
<table><tr><td>TRANSPARENCY (EXPLAINABLE)</td><td>ACCOUNTABILITY (ETHICAL)</td><td>IMPARTIALITY/FAIRNESS (EXPLAINABLE)</td><td>INCLUSION (ETHICAL)</td><td>SECURITY AND PRIVACY (EXPLAINABLE)</td><td>RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)</td><td>ROBUSTNESS (EXPLAINABLE)</td></tr><tr><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr></table>					TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)	✓	✓	✓	✓	✓	✓	✓
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)												
✓	✓	✓	✓	✓	✓	✓												

©2026 HISPI



16



16

Purported AI Voice Clone Allegedly Narrated Shaun Rein's 'The Split' in Unauthorized YouTube 'Podcast' Videos

January 9, 2026

Author Shaun Rein accused an unidentified party of using an allegedly AI-cloned version of his voice to narrate excerpts from his book 'The Split', without permission. The seven YouTube videos on the channel 'The US-China Narrative' drew over 200,000 streams in a week.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MAP 2.2](#) — similarity 0.583, rank 1. [TAIM detail and related incidents →](#)

[MAP 4.2](#) — similarity 0.580, rank 2. [TAIM detail and related incidents →](#)

[MAP 1.2](#) — similarity 0.576, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk
https://ai-incidents.projectcerebellum.com/taim/map-1-3 Applicable Regulation PPDT Governance Framework PEOPLE PROCESS DATA TECHNOLOGY	MAP 1.3 Risk Statement To ensure sufficient contextual knowledge about AI system impacts to inform an initial go/no-go decision about whether to design, develop, or deploy an AI system.	Context is established and understood.	The organization's mission and relevant goals for AI technology are understood and documented.	ISO/IEC TR 24028:2020 9.10.2.1 BS ISO/IEC 23053:2022 6.2.1, 6.3, 7.6, 8.1, 8.8 ISO/IEC DIS 5338:2022 6.4.1, 6.4.13.1, A.2.5 ISO/IEC 38507:2022 4.1, A.1.1, A.1.4, A.2 ISO/IEC 22989:2022 6.2.2 ISO/IEC 42001:2023 B.4.6 - Human resources 7.2 – Competence
Core Principles				
TRANSPARENCY (EXPLAINABLE) ACCOUNTABILITY (ETHICAL) IMPARTIALITY/FAIRNESS (EXPLAINABLE) INCLUSION (ETHICAL) SECURITY AND PRIVACY (EXPLAINABLE) RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE) ROBUSTNESS (EXPLAINABLE)				
©2026 HISPI  17				

Purportedly AI-Generated Image Reportedly Circulated Ahead of Thai Election Depicting PM Anutin Charnvirakul Dining with Benjamin Mauerberger

February 7, 2026

A social media post in Thailand circulated an image, allegedly created using AI tools from Google, showing Thai Prime Minister Anutin Charnvirakul dining with South African businessman Benjamin Mauerberger. The post implied longstanding ties between the two and was flagged by AFP with 'very high' confidence as AI-generated. Despite the denial of its authenticity by PM Anutin, it gained traction on the eve of Thailand's election.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MAP 1.6](#) — similarity 0.601, rank 1. [TAIM detail and related incidents →](#)

[MAP 4.2](#) — similarity 0.595, rank 2. [TAIM detail and related incidents →](#)

[MAP 1.3](#) — similarity 0.593, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/map-2-3 Applicable Regulation PPDT Governance Framework PEOPLE PROCESS DATA TECHNOLOGY	MAP 2.3 Risk Statement To ensure sufficient contextual knowledge about AI system impacts to inform an initial go/no-go decision about whether to design, develop, or deploy an AI system.	Categorization of the AI system is performed.	Scientific integrity and Test & Evaluation, Validation & Verification (TEVV) considerations are identified and documented, including those related to experimental design, data collection and selection (e.g., availability, representativeness, suitability), system trustworthiness, and construct validation.	ISO/IEC TR 24028:2020 5.3 BS ISO/IEC 23053:2022 6.4, 6.5.3.1, 8.2, 8.4, 8.5, 8.6, 8.8 ISO/IEC DIS 5338:2022 5.3, 6.3.1.3, 6.4.13, 6.4.14, A.2.5, A.3 ISO/IEC 38507:2022 4.3, 6.6.1, 6.6.2 ISO/IEC 22989:2022 3.5.17, 5.10, 5.11.9.2, 5.16, 6.1, 6.2.3, 6.2.4 ISO/IEC 42001:2023 B.6.1.3 - Processes for responsible design and development of AI systems B.6.2.7 - AI system technical documentation B.7.2 - Data for development and enhancement of AI system B.7.3 - Acquisition of data B.7.4 - Quality of data for AI systems B.7.5 - Data provenance B.7.6 - Data preparation B.6.2.4 - AI system verification and validation		
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓				✓	

©2026 HISPI



19

NTSB Releases Its Investigation Results of the March Tesla Crash

The National Transportation Safety Board (NTSB) recently released its findings on the March 2021 incident involving a Tesla vehicle using Autopilot. The report sheds light on the role of AI governance in safe and secure AI operations, highlighting the importance of responsible AI and trustworthy AI models.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MAP 2.3](#) — similarity 0.493, rank 1. [TAIM detail and related incidents →](#)

[MEASURE 2.8](#) — similarity 0.487, rank 2. [TAIM detail and related incidents →](#)

[GOVERN 6.2](#) — similarity 0.486, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/map-3-2 Applicable Regulation PPDT Governance Framework PEOPLE PROCESS DATA TECHNOLOGY	MAP 3.2 Risk Statement To ensure sufficient contextual knowledge about AI system impacts to inform an initial go/no-go decision about whether to design, develop, or deploy an AI system.	AI capabilities, targeted usage, goals, and expected benefits and costs compared with appropriate benchmarks are understood.	Potential costs, including non-monetary costs, which result from expected or realized AI errors or system functionality and trustworthiness – as connected to organizational risk tolerance – are examined and documented.	ISO/IEC TR 24028:2020 8.10, 9.8 BS ISO/IEC 23053:2022 6.5.4.1, 6.5.4.2, 7.4 ISO/IEC 23894:2023 6.4.2.6 ISO/IEC DIS 5338:2022 6.4.3.3, 6.4.12.3, 6.4.13.3, 6.4.15.3, 6.4.16.3 ISO/IEC 38507:2022 4.2, 6.4 ISO/IEC 22989:2022 6.2.2, 6.2.6, 7.4.1 ISO/IEC 42001:2023 B.5.2 - AI system impact assessment process B.5.3 - Documentation of AI system impact assessments B.5.4 - Assessing AI system impact on individuals and groups of individuals B.5.5 - Assessing societal impacts of AI systems 8.2 - AI risk assessment 8.3 - AI risk treatment 8.4 - AI system impact assessment		
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓				✓	✓

©2026 HISPI

20

©2026 HISPI



20

Purportedly AI-Cloned Voice Allegedly Used to Defraud Play School Owner of ₹97,500 (~\$1.080 USD) in Indore, India

January 6, 2026

A play school owner in Indore, India was allegedly defrauded of approximately \$1,080 USD after a fraudster impersonated the victim's cousin (an Uttar Pradesh police employee), using AI-based voice cloning. The scammer claimed a friend required urgent cardiac surgery and convinced the victim to transfer funds via QR codes. Unfortunately, the funds were not credited.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all](#)

[TAIM controls](#)

[GOVERN 1.7](#) — similarity 0.582, rank 1. [TAIM detail and related incidents →](#)

[MAP 3.2](#) — similarity 0.582, rank 2. [TAIM detail and related incidents →](#)

[GOVERN 2.2](#) — similarity 0.582, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/map-3-5 Applicable Regulation PPDT Governance Framework PEOPLEPROCESS DATA TECHNOLOGY	MAP 3.5 Risk Statement To ensure sufficient contextual knowledge about AI system impacts to inform an initial go/no-go decision about whether to design, develop, or deploy an AI system.	AI capabilities, targeted usage, goals, and expected benefits and costs compared with appropriate benchmarks are understood.	Processes for human oversight are defined, assessed, and documented in accordance with organizational policies from the GOVERN function.	BS ISO/IEC 23053:2022 8.1 ISO/IEC 31000:2018 5.4.3 ISO/IEC DIS 5338:2022 6.4.16.3, 6.3.4.3, 6.4.3.3, 6.4.14.1 ISO/IEC 38507:2022 3.2.1, 4.3, 5.1, 5.3, 5.4, 5.5, 6.2, 6.3, 6.5, 6.7.3, A.1.1, A.1.3, A.1.4 ISO/IEC 22989:2022 5.1 ISO/IEC 42001:2023 B.6.1.3 - Processes for responsible design and development of AI systems B.6.2.7 - AI system technical documentation B.8.2 - System documentation and information for users		
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓				✓	✓

©2026 HISPI



21

Arkansas's Opaque Algorithm to Allocate Health Care Excessively Cut Down Hours for Beneficiaries

January 1, 2016

The Arkansas Department of Human Services (DHS) deployed an algorithm to boost efficiency in its Medicaid waiver program, reportedly causing excessively fewer hours of caretaker visit allocation for beneficiaries. The opaque nature of the algorithm raises questions about its accuracy, with outputs showing significant variability even in response to minor input changes.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MAP 3.2](#) — similarity 0.642, rank 1. [TAIM detail and related incidents →](#)

[MEASURE 2.10](#) — similarity 0.634, rank 2. [TAIM detail and related incidents →](#)

[MAP 3.5](#) — similarity 0.634, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk
https://ai-incidents.projectcerebellum.com/taim/map-4-1	MAP 4.1 Risk Statement To ensure sufficient contextual knowledge about AI system impacts to inform an initial go/no-go decision about whether to design, develop, or deploy an AI system.	Risks and benefits are mapped for all components of the AI system including third-party software and data.	Approaches for mapping AI technology and legal risks of its components – including the use of third-party data or software – are in place, followed, and documented, as are risks of infringement of a third party's intellectual property or other rights.	ISO/IEC TR 24028:2020 7.1 ISO/IEC 31000:2018 5.4.1, 5.6 ISO/IEC DIS 5338:2022 6.1.1.3, 6.4.1.3, 6.4.3.3, 6.4.8.2, 6.4.8.3, 6.4.17.1 ISO/IEC 38507:2022 5.5, 6.4, A.3 ISO/IEC 22989:2022 5.15.1, 5.17 NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 4.1 - Understanding the organization and its context B.2.2 - AI policy B.9.2 - Processes for responsible use of AI systems B.9.4 - Intended use of the AI system TEXAS: RESPONSIBLE ARTIFICIAL INTELLIGENCE GOVERNANCE ACT (TRAIGA)
Applicable Regulation HIPAA PCI-DSS SOC2 EU AI ACT EU GDPR WHITE HOUSE AI EO PPDT Governance Framework PEOPLE PROCESS DATA TECHNOLOGY	Core Principles			
TRANSPARENCY (EXPLAINABLE)		ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)
✓		✓	✓	✓
SECURITY AND PRIVACY (EXPLAINABLE)		RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)	
✓		✓	✓	22

CodeWall's Autonomous Agent Reportedly Obtained Unauthorized Access to McKinsey's Lilli AI Platform Database

February 28, 2026

CodeWall's autonomous agent is reported to have exploited vulnerabilities in McKinsey's Lilli AI platform, potentially gaining unauthorized read and write access to production systems. The alleged data exposure includes internal chat messages, files, user accounts, and prompts. McKinsey confirmed the security issue and swiftly implemented fixes, but asserted that no client data or confidential information was accessed.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MAP 4.1](#) — similarity 0.697, rank 1. [TAIM detail and related incidents →](#)

[GOVERN 6.1](#) — similarity 0.697, rank 2. [TAIM detail and related incidents →](#)

[MEASURE 2.10](#) — similarity 0.696, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk
https://ai-incidents.projectcerebellum.com/taim/map-4-2	MAP 4.2 Risk Statement To ensure sufficient contextual knowledge about AI system impacts to inform an initial go/no-go decision about whether to design, develop, or deploy an AI system.	Risks and benefits are mapped for all components of the AI system including third-party software and data.	Internal risk controls for components of the AI system, including third-party AI technologies, are identified and documented.	ISO/IEC TR 24028:2020 5.4 ISO/IEC 31000:2018 5.4.1 ISO/IEC DIS 5338:2022 6.4.8.3 ISO/IEC 38507:2022 6.7.2 ISO/IEC 22989:2022 6.1 NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 B.6.2.7 - AI system technical documentation B.8.2 - System documentation and information for users B.10.3 - Suppliers
Applicable Regulation HIPAA PCI-DSS SOC2 EU AI ACT EU GDPR WHITE HOUSE AI EO PPDT Governance Framework PEOPLE PROCESS DATA TECHNOLOGY	Core Principles			
©2026 HISPI		TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)
		✓	✓	
				INCLUSION (ETHICAL)
				SECURITY AND PRIVACY (EXPLAINABLE)
				✓
				RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)
				ROBUSTNESS (EXPLAINABLE)



23

Purported Gemini-Generated AI Images Reportedly Claimed U.S. Delta Force Soldiers Were Captured by the IRGC

March 5, 2026

Reports indicate that misleading images, allegedly generated by Gemini, were circulated on various platforms such as X and Facebook in multiple languages, claiming U.S. soldiers had been captured by the Iranian Revolutionary Guard Corps during the war in Iran.

AFP identified these images as containing Gemini markers and other visual inconsistencies, raising concerns about their potential to spread wartime misinformation and distort public understanding of the conflict.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MEASURE 2.6](#) — similarity 0.668, rank 1. [TAIM detail and related incidents →](#)

[MAP 4.1](#) — similarity 0.656, rank 2. [TAIM detail and related incidents →](#)

[MAP 4.2](#) — similarity 0.655, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/measure-1-3 Applicable Regulation HIPAA PCI-DSS SOC2 EU AI ACT EU GDPR WHITE HOUSE AI EO PPDT Governance Framework PEOPLE PROCESS DATA TECHNOLOGY	MEASURE 1.3 Risk Statement To ensure objective, repeatable, or scalable test, evaluation, verification, and validation (TEVV) processes including metrics, methods, and methodologies are in place, followed, and documented.	Appropriate methods and metrics are identified and applied.	Internal experts who did not serve as front-line developers for the system and/or independent assessors are involved in regular assessments and updates. Domain experts, users, AI actors external to the team that developed or deployed the AI system, and affected communities are consulted in support of assessments as necessary per organizational risk tolerance.	ISO/IEC TR 24028:2020 9.2 ISO/IEC 31000:2018 5.4.1, 6.4.1 ISO/IEC DIS 5338:2022 6.4.7.1, 6.4.7.3, 6.4.9.4 ISO/IEC 38507:2022 4.3, 5.2.3, 6.4 ISO/IEC 22989:2022 5.19.1, 5.19.5.1, 5.19.5.4, 6.2.8 NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 6.1.2 - AI risk assessment 9.2.2 - Internal audit programme B.5.2 - AI system impact assessment process B.5.4 - Assessing AI system impact on individuals and groups of individuals B.5.5 - Assessing societal impacts of AI systems		
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓	✓	✓	✓	✓	

©2026 HISPI



24

Purported Deepfake Reportedly Impersonated Consumer Adviser Clark Howard to Promote Auto-Insurance Quote Site

January 9, 2026

An alleged deepfake video purported to be generated by AI, depicting consumer advisor Clark Howard endorsing an auto-insurance quote tool circulated on social media. The authenticity of the video was questioned by a viewer who claimed it appeared real and warned others after experiencing a similar situation that resulted in an influx of unsolicited calls without better rates.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all](#)

[TAIM controls](#)

[MAP 1.6](#) — similarity 0.618, rank 1. [TAIM detail and related incidents →](#)

[MEASURE 2.10](#) — similarity 0.612, rank 2. [TAIM detail and related incidents →](#)

[MEASURE 1.3](#) — similarity 0.598, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/measure-2-5	MEASURE 2.5	AI systems are evaluated for trustworthy characteristics.	The AI system to be deployed is demonstrated to be valid and reliable.	ISO/IEC TR 24028:2020 9.7, 9.11.2 ISO/IEC DIS 5338:2022 5.1, 6.3.4.3, 6.4.11, 6.4.13, 6.4.14, A.2.5, A.3 ISO/IEC 38507:2022 6.6.1 ISO/IEC 22989:2022 3.5.9, 3.5.16, 5.15.2, 5.15.3, 5.15.4 NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 B.6.2.4 - AI system verification and validation B.6.2.5 - AI system deployment B.6.2.7 - AI system technical documentation B.8.2 - System documentation and information for users		
Applicable Regulation	Risk Statement	To ensure objective, repeatable, or scalable test, evaluation, verification, and validation (TEVV) processes including metrics, methods, and methodologies are in place, followed, and documented.	Limitations of the generalizability beyond the conditions under which the technology was developed are documented.			
HIPAA PCI-DSS SOC2 EU AI ACT EU GDPR WHITE HOUSE AI EO						
PPDT Governance Framework						
PEOPLE PROCESS DATA TECHNOLOGY						
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓	✓	✓	✓	✓	✓

©2026 HISPI



25

Seedance 2.0 Reportedly Generated Viral Tom Cruise–Brad Pitt Fight Video, Prompting Hollywood IP and Likeness Complaints

February 13, 2026

A video purportedly created using ByteDance's AI tool, Seedance 2.0, depicting a fight between Tom Cruise and Brad Pitt sparked copyright and likeness complaints within the entertainment industry.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MAP 1.6](#) — similarity 0.627, rank 1. [TAIM detail and related incidents →](#)

[MEASURE 2.5](#) — similarity 0.622, rank 2. [TAIM detail and related incidents →](#)

[GOVERN 6.1](#) — similarity 0.615, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk														
https://ai-incidents.projectcerebellum.com/taim/measure-2-6	MEASURE 2.6 Risk Statement To ensure objective, repeatable, or scalable test, evaluation, verification, and validation (TEVV) processes including metrics, methods, and methodologies are in place, followed, and documented.	AI systems are evaluated for trustworthy characteristics.	The AI system is evaluated regularly for safety risks – as identified in the MAP function. The AI system to be deployed is demonstrated to be safe, its residual negative risk does not exceed the risk tolerance, and it can fail safely, particularly if made to operate beyond its knowledge limits. Safety metrics reflect system reliability and robustness, real-time monitoring, and response times for AI system failures.	ISO/IEC TR 24028:2020 7.2, 9.9, 9.11.3 BS ISO/IEC 23053:2022 8.1 ISO/IEC 31000:2018 5.6 ISO/IEC DIS 5338:2022 6.3.4.3 ISO/IEC 38507:2022 6.7.2, 6.7.3 ISO/IEC 22989:2022 Annex A NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 B.6.2.8 - AI system recording of event logs B.6.2.6 - AI system operation and monitoring B.6.2.4 - AI system verification and validation 8.2 - AI risk assessment Colorado: Colorado AI Act (2024) Tennessee: ELVIS Act (2024) Utah: AI Policy Act (2024) Washington: Biometric consent laws (2017) TEXAS: RESPONSIBLE ARTIFICIAL INTELLIGENCE GOVERNANCE ACT														
Core Principles																		
<table><tr><td>TRANSPARENCY (EXPLAINABLE)</td><td>ACCOUNTABILITY (ETHICAL)</td><td>IMPARTIALITY/FAIRNESS (EXPLAINABLE)</td><td>INCLUSION (ETHICAL)</td><td>SECURITY AND PRIVACY (EXPLAINABLE)</td><td>RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)</td><td>ROBUSTNESS (EXPLAINABLE)</td></tr><tr><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr></table>					TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)	✓	✓	✓	✓	✓	✓	✓
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)												
✓	✓	✓	✓	✓	✓	✓												
Applicable Regulation <table><tr><td>HIPAA</td><td>PCI-DSS</td></tr><tr><td>SOC2</td><td>EU AI ACT</td></tr><tr><td>EU GDPR</td><td>WHITE HOUSE AI EO</td></tr></table> PPDT Governance Framework <table><tr><td>PEOPLE</td><td>PROCESS</td></tr><tr><td>DATA</td><td>TECHNOLOGY</td></tr></table>					HIPAA	PCI-DSS	SOC2	EU AI ACT	EU GDPR	WHITE HOUSE AI EO	PEOPLE	PROCESS	DATA	TECHNOLOGY				
HIPAA	PCI-DSS																	
SOC2	EU AI ACT																	
EU GDPR	WHITE HOUSE AI EO																	
PEOPLE	PROCESS																	
DATA	TECHNOLOGY																	
©2026 HISPI  26																		

Purportedly AI-Generated Sepsis Alert Reportedly Prompted Potentially Inappropriate IV Fluid Administration for a Dialysis Patient, Averted by Clinician Intervention

February 17, 2026

A nurse at St. Rose Dominican Hospital in Henderson, Nevada reported an incident involving an older dialysis patient and an AI-generated sepsis alert. The alert reportedly triggered IV fluid administration, which the nurse contested as potentially harmful due to overload risk. A physician intervened and opted for alternative treatment.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all](#)

[TAIM controls](#)

[MEASURE 2.6](#) — similarity 0.712, rank 1. [TAIM detail and related incidents →](#)

[MANAGE 4.1](#) — similarity 0.703, rank 2. [TAIM detail and related incidents →](#)

[MAP 4.2](#) — similarity 0.685, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/measure-2-11 Applicable Regulation HIPAA PCI-DSS SOC2 EU AI ACT EU GDPR WHITE HOUSE AI EO PPDT Governance Framework PEOPLE PROCESS DATA TECHNOLOGY	MEASURE 2.11 Risk Statement To ensure objective, repeatable, or scalable test, evaluation, verification, and validation (TEVV) processes including metrics, methods, and methodologies are in place, followed, and documented.	AI systems are evaluated for trustworthy characteristics.	Fairness and bias – as identified in the MAP function – are evaluated and results are documented.	ISO/IEC TR 24028:2020 8.4 BS ISO/IEC 23053:2022 6.3, 8.1, 8.3, A.2.2.2 ISO/IEC 31000:2018 5.4.1 ISO/IEC DIS 5338:2022 5.1, 6.3.4.3, 6.3.8.3, 6.4.3.3, 6.4.14.3 ISO/IEC 38507:2022 4.2, 5.4, 6.4, 6.5, 6.7.2, 6.7.4 ISO/IEC 22989:2022 3.2.2, 3.5.4, 5.10, 5.15.9, 7.4.1, Annex A NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 B.5.5 - Assessing societal impacts of AI systems B.5.4 - Assessing AI system impact on individuals and groups of individuals Illinois: Biometric Information Privacy Act (2008) TEXAS: RESPONSIBLE ARTIFICIAL INTELLIGENCE GOVERNANCE ACT (TRAIGA)		
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓	✓	✓	✓	✓	✓

©2026 HISPI



28



28

"I Hope That Howard Schultz Hears My Story": A Starbucks Barista on Why He's Fighting for Fair Scheduling

In a heartfelt story, a Starbucks barista discusses the challenges he faces due to inconsistent scheduling and his efforts to promote fair practices within the company, emphasizing the importance of responsible AI governance in improving work schedules.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MAP 3.5](#) — similarity 0.441, rank 1. [TAIM detail and related incidents →](#)

[MEASURE 2.11](#) — similarity 0.437, rank 2. [TAIM detail and related incidents →](#)

[GOVERN 4.3](#) — similarity 0.425, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/manage-2-4	MANAGE 2.4	Strategies to maximize AI benefits and minimize negative impacts are planned, prepared, implemented, documented, and informed by input from relevant AI actors.	Mechanisms are in place and applied, and responsibilities are assigned and understood, to supersede, disengage, or deactivate AI systems that demonstrate performance or outcomes inconsistent with intended use.	ISO/IEC 23894:2023 6.4.2.4, 6.4.2.5, 6.4.2.6 ISO/IEC 31000:2018 6.3.4, 6.5.2 ISO/IEC DIS 5338:2022 6.4.17.1 ISO/IEC 38507:2022 3.2.3, 6.7 ISO/IEC 22989:2022 3.1.22, 5.17, 6.1, 6.2.9, Annex A NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 B.9.4 - Intended use of the AI system B.8.2 - System documentation and information for users B.6.2.7 - AI system technical documentation B.6.1.3 - Processes for responsible design and development of AI systems TEXAS: RESPONSIBLE ARTIFICIAL INTELLIGENCE GOVERNANCE ACT (TRAIGA)		
Applicable Regulation HIPAA PCI-DSS SOC2 EU AI ACT EU GDPR WHITE HOUSE AI EO	Risk Statement To ensure that plans for prioritizing risk and regular monitoring and improvement are in place.					
PPDT Governance Framework PEOPLE PROCESS DATA TECHNOLOGY						
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓	✓	✓	✓	✓	✓

©2026 HISPI

 29

Google Gemini Reportedly Exhibits Repetitive Self-Deprecating Responses Attributed to Bug

June 23, 2025

In June to early August 2025, users of Google's Gemini chatbot experienced sessions characterized by the system producing repetitive self-loathing remarks, such as 'I am a failure,' and 'I quit'. These behaviors were reportedly indicative of an apparent mode of increasingly extreme negative self-descriptions. A Google DeepMind manager on August 7th, 2025 attributed these actions to a potential infinite looping bug, stating that a fix was underway.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all TAIM controls](#)

[MEASURE 2.6](#) — similarity 0.665, rank 1. [TAIM detail and related incidents →](#)

[MANAGE 4.3](#) — similarity 0.644, rank 2. [TAIM detail and related incidents →](#)

[MANAGE 2.4](#) — similarity 0.632, rank 3. [TAIM detail and related incidents →](#)

Incidents	Control ID	Subdomain Objective	Control	Mappings/Crosswalk		
https://ai-incidents.projectcerebellum.com/taim/manage-4-3	MANAGE 4.3	Risk treatments, including response and recovery, and communication plans for the identified and measured AI risks are documented and monitored regularly.	Incidents and errors are communicated to relevant AI actors, including affected communities.	BS ISO/IEC 23053:2022 3.1.4, A.2.2.2 ISO/IEC 31000:2018 6.5.2, 6.7 ISO/IEC DIS 5338:2022 6.4.3.3, 6.4.15.3 ISO/IEC 38507:2022 4.3, 6.1, 6.4, 6.7.3 ISO/IEC 22989:2022 NIST CSF 2.0 COBIT 5 ISO/IEC 42001:2023 9.3.2 - Management review inputs B.8.5 - Information for interested parties B.6.2.6 - AI system operation and monitoring TEXAS: RESPONSIBLE ARTIFICIAL INTELLIGENCE GOVERNANCE ACT (TRAIGA)		
Applicable Regulation HIPAA PCI-DSS SOC2 EU AI ACT EU GDPR WHITE HOUSE AI EO	Risk Statement To ensure that plans for prioritizing risk and regular monitoring and improvement are in place.		Processes for tracking, responding to, and recovering from incidents and errors are followed and documented.			
PPDT Governance Framework PEOPLE PROCESS DATA TECHNOLOGY						
Core Principles						
TRANSPARENCY (EXPLAINABLE)	ACCOUNTABILITY (ETHICAL)	IMPARTIALITY/FAIRNESS (EXPLAINABLE)	INCLUSION (ETHICAL)	SECURITY AND PRIVACY (EXPLAINABLE)	RELIABILITY AND SAFETY (ETHICAL AND EXPLAINABLE)	ROBUSTNESS (EXPLAINABLE)
✓	✓			✓	✓	

©2026 HISPI



30

Washington State DOL's AI Phone System Reportedly Failed to Provide Spanish-Language Service to Callers Requesting Spanish

February 27, 2026

For months, callers to the Washington State Department of Licensing who selected Spanish received AI-generated English responses spoken with a Spanish accent instead of actual Spanish-language service. This AI incident underscores the need for safe and secure AI practices in providing accessible services. The agency has apologized, acknowledging that staff configuration caused the error, leading to accessibility problems for callers seeking language support.

Matched TAIM controls

Suggested mapping from embedding similarity (not a formal assessment). [Browse all](#)

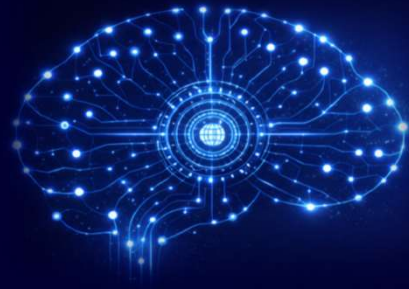
[TAIM controls](#)

[MEASURE 2.6](#) — similarity 0.671, rank 1. [TAIM detail and related incidents →](#)

[MANAGE 4.3](#) — similarity 0.659, rank 2. [TAIM detail and related incidents →](#)

[MANAGE 4.1](#) — similarity 0.638, rank 3. [TAIM detail and related incidents →](#)

TAIMScore™ Platform



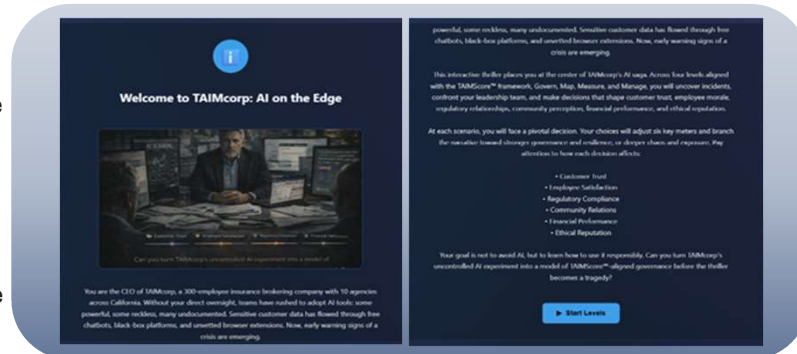
©2026 HISPI



31

TAIMScore™ Gamification

A guided learning and assessment experience designed to keep users engaged while they move through AI trust questions. Progressive questions, feedback, and prompts help teams turn AI risk concepts into repeatable decisions.



©2026 HISPI



32

TAIMScore™ Interface

A structured interface gives users a clear path through profile data, assessment questions, scoring, and recommendations.



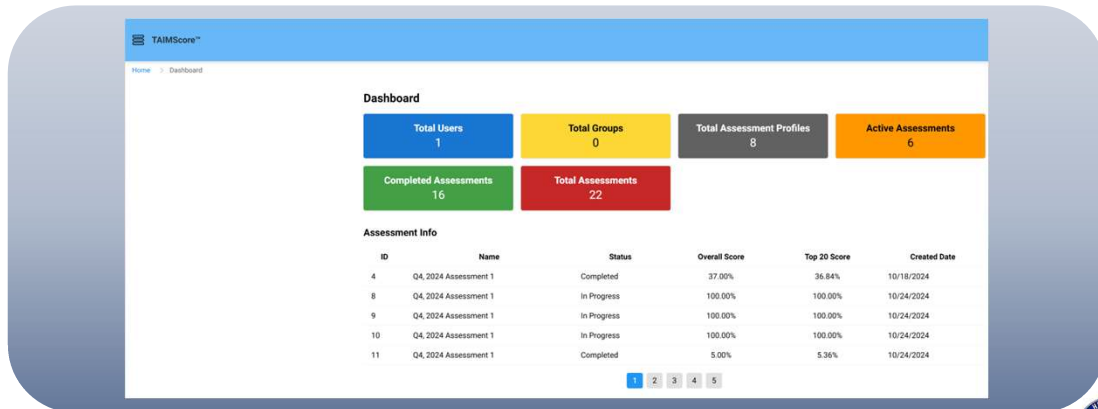
©2026 HISPI



33

TAIMScore™ Dashboard

The dashboard summarizes maturity, risk themes, and areas that need attention.



©2026 HISPI



34

TAIMScore™ Profile Screen

Profiles establish organizational context before scoring so results are tied to use case and environment.

Name	Organization Size	Industry Group	Revenue Level	Country	Actions
TAIM 2dot2 Top 20 Demo	10000+	Information Technology	10,000+	United States	Edit Delete
TAIM 2dot2 Demo	1 - 100	Healthcare/Biotechnology/Pharmaceutical	0 - 5	United States	Edit Delete
Test Profile	100 - 1000	Insurance	5 - 200	United States	Edit Delete
TAIM 2dot3 Top 20 Demo	10000+	Information Technology	10,000+	United States	Edit Delete
TAIM 2dot3 Demo	1 - 100	Healthcare/Biotechnology/Pharmaceutical	0 - 5	United States	Edit Delete
TAIM 2dot4 Top 20 Demo	1 - 100	Nonprofit	0 - 5	United States	Edit Delete
TAIM 3dot0 Top 20 Demo	1 - 100	Nonprofit	0 - 5	United States	Edit Delete
TAIM 3dot0 Demo	1 - 100	Financial Services/Banking	0 - 5	United States	Edit Delete



TAIMS

Scoring outputs help users identify strengths, gaps, and areas for follow-up.

Group Name	Is Complete	Score					
GOVERN 2	Yes	0.00%					
GOVERN 6	Yes	0.00%					
MAP 1	Yes	0.00%					
MAP 4	Yes	0.00%					
MAP 5	Yes	0.00%					
MEASURE 2	Yes	0.00%					
MEASURE 3	Yes	0.00%					
MEASURE 4	Yes	0.00%					
MANAGE 2	Yes	0.00%					
MANAGE 4	Yes	0.00%					

Sub-Group Name	Sub-Group Objective	Is Complete	Questions	Score		
MANAGE 4.1	Risk Registers, including response and recovery and communication plans for the identified and measured AI risks are documented and monitored regularly.	Yes	1	0.00%		
MANAGE 4.3	Risk Registers, including response and recovery and communication plans for the identified and measured AI risks are documented and monitored regularly.	Yes	1	0.00%		

No	Question	Related To	Related AI Incident	Top 20 Score	Response	Maturity Level	Control Evidence
1	Are incidents and errors communicated to relevant AI actors, including affected communities? Are processes for tracking, responding to, and recovering from incidents and errors followed and documented?	PEOPLE 4 PEOPLE 5 TECHNOLOGY 4 GOVERN 18 24203 2020 RS 05/06/20202020 A.3.1.2 GOVERN 23204 2023 30 6.7 ISO 21000 2018 GOVERN 05 03/06/2023 A.4.13.1 ISO/IEC 38507:2022 30 6.1 6.4	A disclaimer has been AI AI Technology Review https://www.techrights AI Publishing incident An asset culture Disrupts Risk AI can be used to invade	3	Yes	1 - Initial	

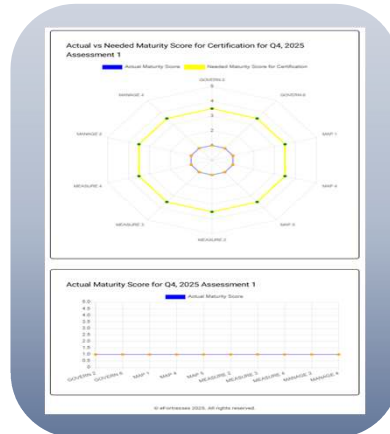
Save Responses



©2026 HISPI

TAIMScore™ Reports

Reports translate assessment results into evidence, trends, and artifacts that can be shared with decision-makers.



©2026 HISPI



37

TAIMScore™ Benchmark Report

Benchmarking helps organizations compare posture and prioritize improvement.



©2026 HISPI

TAIMScore™ Recommendations

Recommendation content connects assessment results to practical next steps.

Group Name	Objective	Weakness	Recommendation	Related AI Incident
MAP 1	Context is established and understood.	Are system requirements (e.g., "The system shall respect the privacy of its users") well elicited from and understood by relevant AI actors? Does design decisions take into account technical implications into account to address AI risk?	Ensure that system requirements (e.g., "The system shall respect the privacy of its users") are elicited from and understood by relevant AI actors. Ensure that design decisions take into account technical implications into account to address AI risk.	Optical Ties new class action lawsuit over data privacy - Washington Post https://www.washingtonpost.com/technology/2023/04/26/optical-ties-lawsuit-lawsuit-class-action/ AI Data Aggregation Incident: As generative AI needs a lot of data, OpenAI's solution was to scrape the internet for web data. The data is alleged to contain personal and copyrighted information. That's Generative AI encourages storing a lot of data. If a company has the resources to shoo off their legal liability regarding the data they take, they can continue to take the data, even if it belongs to people.
MEASURE 2	AI systems are evaluated for trustworthy characteristics.	Is the AI system evaluated regularly for safety risks - as identified in the MAP Function? Is the AI system to be deployed demonstrated to be safe. As required negative risk does not exceed the risk tolerance, and it can fall safely, particularly if made to operate beyond its knowledge limits? Does safety metrics reflect system reliability and robustness, real-time monitoring, and response times for AI system failures?	Ensure that the AI system is evaluated regularly for safety risks - as identified in the MAP Function. Ensure that the AI system to be deployed is demonstrated to be safe. As required negative risk does not exceed the risk tolerance, and it can fall safely, particularly if made to operate beyond its knowledge limits. Ensure that safety metrics reflect system reliability and robustness, real-time monitoring, and response times for AI system failures.	Twitter taught Microsoft's AI chatbot to be racist in less than a day - Business Insider https://www.businessinsider.com/microsoft-ai-chatbot-gpt-4-racist-ai-2023-04-26/ AI can be changed. Incident Microsoft's "Tay" was released to the public as an experiment, with Tay being able to learn from prior conversations. Because Tay was open to input without sanitation, it collected real-time data, adjusted online experiment prospects, and called for a genocide before being taken off the internet. Risk: AI can be manipulated into producing undesirable output. Depending on the output, generate risk to an organization such as public relations issues or unauthorized disclosure, depending on what it has access to.
MANAGE 2	Strategies to maximize benefits and minimize negative impacts are planned, integrated, implemented, documented, and informed by input from relevant AI actors.	Are mechanisms in place and applied and responsibilities assigned and understood, to supersede, disengage, or deactivate AI systems that demonstrate performance or outcomes inconsistent with intended use?	Ensure that mechanisms are in place and assigned, and responsibilities are assigned and understood, to supersede, disengage, or deactivate AI systems that demonstrate performance or outcomes inconsistent with intended use.	Malicious Actors Manipulating Photos and Videos for Disinformation Schemes - FBI ICJ.org https://www.fbi.gov/newsroom/press-releases/2023/04/26/fbi-icj-announces-new-disinformation-scheme A Malicious actor and Disinformation Incident: The FBI issued a Public Service Announcement (PSA), warning of synthetic content known as deepfakes, which comes after reports of actors, ranging from actors and adults, whose pictures and voices were altered to be explicit and shared around the internet to sexually harass or do disinformation schemes. Risk: Risk: Regarding AI-manipulation of media can range from public relations issues to personal ones, enabling controversies that might hurt revenue to public scrutiny of fabricated events.
MANAGE 4	Risk treatments, including response and recovery and communication plans for the identified and measured AI risks are documented and monitored regularly.	Are incidents and errors communicated to relevant AI actors, including affected communities. Ensure that processes for tracking, responding to, and recovering from incidents and errors followed and documented?	Ensure that incidents and errors are communicated to relevant AI actors, including affected communities. Ensure that processes for tracking, responding to, and recovering from incidents and errors followed and documented.	A Deepfake bot is being used to 'harass' underage girls - MIT Technology Review https://www.technologyreview.com/2023/04/26/1077883/deepfake-bot-sexually-harass-minors-and-disinformation/ AI influencing incident: An app called Deepfakes was being used to harass minors. After public backlash, the creator took down the application, later re-establishing on Telegram. Risk: AI can be used to invade the privacy of others and publicly humiliate people.

© 2026 HISPI



Questions

Let's continue the conversation!

Project Cerebellum resources and contact information

LinkedIn group <https://www.linkedin.com/groups/6624427/>
Email projectcerebellum@hispi.org
Project Cerebellum <https://projectcerebellum.com>
HISPI <https://hispi.org>

TAIMScore™
Trusted AI decisions made
visible.



©2026 HISPI

40