# OECD AI Transparency Report

**Organization:** SUHUL LLC (USA) **Reporting Period:** Q4 2026 **Published:** November 10, 2026

## Section 0 – Your role in the AI system lifecycle

**Responding as (please select all that apply):**

- Application developer (/provider)
- Deployer

> **If any of the fields above are selected, please indicate your organization size:**
>
> Small enterprise: 10-49 employees

## Section 1 – Risk Identification and Evaluation

**1. How does your organization define and/or classify different types of risks related to AI?[1]**

- Maintain an internal taxonomy, classification system, and/or set of risk area definitions
- Align with specific external frameworks, standards, reports, and/or guidance documents such as taxonomies, classification systems, and/or sets of definitions for risk areas (e.g. the NIST AI Risk Management Framework, the ISO/IEC 42001, the OECD AI Principles, the EU General-Purpose AI Code of Practice, the International AI Safety Report). (please specify): NIST AI RMF 1.0 and NIST AI 600-1 (Generative AI Profile); OECD AI Principles; CSA AI Controls Matrix (AICM v1.1); our information security management system policy.

> **How does your organization define thresholds at which severe risks, unless adequately mitigated, may be considered "unreasonable"? Please describe the framework(s) most relevant to your organization's role and context. If your organisation does not use formal thresholds, you may describe instead the policies, decision criteria, or governance processes you apply when assessing severe AI risks. You may also describe: • Specific examples of contexts in which your models or systems could, in your assessment, pose unreasonable or intolerable risk; • How mitigations are graduated by capability or propensity level; • How your organisation considers the specificities of different AI technologies, such as generative AI and agentic AI, including any specific risk assessment or governance mechanisms applied to agentic AI capabilities.**
>
> SUHUL does not train or fine-tune foundation models; we build business applications on top of third-party models (Anthropic Claude, Cloudflare Workers AI, and Google Vertex AI in the Taiwan region). Our risk thresholds therefore apply to what an AI feature can read and do inside a customer's workspace.

> We treat the following as unreasonable unless mitigated: (1) customer content leaving the processing locations the customer has agreed to; (2) an AI agent completing an irreversible or financial action without a human decision; (3) instructions hidden in documents, tickets or external content being followed by the model; (4) personal identifiers being sent to an external model provider when the customer's policy requires de-identification.
> Mitigations are graduated by the capability of each feature: read-only summarisation and classification run with logging only; features that draft content are labelled as AI suggestions; agentic features that call tools are registered with a permission and risk level, and actions that change records or contact third parties require explicit human confirmation bound to the exact parameters; high-risk actions such as refunds, rate publication and financial period closing are technically blocked for AI agents and can only be performed by a person.

**2. Both before and after deployment, what practices does your organization use to identify and evaluate vulnerabilities, incidents, emerging risks and misuse throughout the AI system lifecycle?**

- Adversarial testing to detect technical vulnerabilities
- Internal benchmarks, standard metrics (e.g., frequency, severity scores)
- Qualitative assessments (e.g., expert judgment, narrative scoring)
- Impact assessments such as personal data protection assessment and fundamental rights impact assessments
- Vulnerability reports
- Incident reports, including reports shared by other organizations
- Automated monitoring / remote status reporting (e.g., telemetry pipelines, automated alerts when risk thresholds are breached)
- Specific risk identification practices for agentic AI capabilities, e.g., autonomous action-taking, tool use, multi-step planning (please specify): Each AI agent tool is registered with a permission level and risk level. Tool actions that change records or contact third parties require explicit human confirmation, bound to the exact parameters approved. Tool results, attachments and external content are screened for prompt injection and treated as untrusted data. AI agents are technically blocked from high-risk actions such as refunds, rate publication and financial period closing.

> **Please briefly describe your organization’s decision-making process for deploying an AI model or system including: who has decision authority; what evidence and criteria are required (e.g., performance, safety/security, robustness, fairness, privacy, compliance, independent external testing, evaluation, validation and verification (TEVV) or assurance); contextual caveats or limitations required in interpreting risk evaluation metric results; what approvals/sign-offs are needed; and what safeguards (e.g., mitigations, monitoring, or phased release) must be in place to move into deployment.**
>
> Decision authority sits with the Governor and the R&D lead. A new AI feature or a change in the model used for a feature is deployed only when: the data it reads and the actions it can take are documented; it goes through the single AI egress policy (customer content stays on Cloudflare Workers AI unless the customer opts in to other providers; Taiwan-only customers are processed only in the Taiwan region); usage is metered through the common entitlement check; prompt-injection handling and human confirmation are in place for any tool actions; and automated type checks and tests pass in a clean build.

Every production release goes through a deployment gate that runs these checks, uploads the new version with 0% traffic, and requires a sign-off by the Governor before traffic is switched. Evaluation results are interpreted with the caveat that model behaviour can vary between versions of third-party models, so behaviour is monitored after release and features can be rolled back.

**Following deployment, how do you use incident reports — including those shared by other organizations — to identify emerging issues, update risk assessments, and trigger corrective actions.**

Each AI call is logged with the feature, model, customer workspace, operator and cost, and AI outcomes (success, failure, user corrections) are recorded. Failures and degradations are written to a central system-health event log that drives alerts. When an incident or a new attack pattern is found — including issues published by model providers, security researchers or other organisations — we reproduce it, add a regression test, update the affected safeguard (for example, the prompt-injection screening rules were extended after a real case where instructions were hidden inside a store name), and review other features that share the same pattern. Security issues can be reported through our public vulnerability disclosure policy (https://suhul.us/.well-known/security.txt). A formal AI incident register is being added to our security incident process.

**3. Does your organization have mechanisms to receive and/or share reports of risks, incidents or vulnerabilities by/with third parties?**

Yes, receiving only

**If yes, please indicate which stakeholders reporting mechanisms may be accessible to, recognising specifics will depend on the nature of the report and risk level. For each, indicate whether access is on a standing basis (Always), activated based on context (Ad hoc/as relevant), or not currently accessible.**

| | Always | Ad hoc / as relevant | Not currently accessible |
|---|---|---|---|
| Relevant employees (security, legal/compliance, etc.) | □ | | |
| Public authorities (e.g., public reporting offices for security incidents or data breaches) | | □ | |
| Supervisory authorities (e.g., data protection authorities) | | | □ |
| Industry partners or industry fora | | □ | |
| Users (end users, enterprise customers, developers using APIs) | □ | | |
| Business customers / deploying organizations (if different from “users”) | □ | | |
| Supply-chain partners (cloud, data providers, etc.) | | □ | |
| Academia / researchers | | | □ |
| Independent auditors | | | □ |

| | Always | Ad hoc / as relevant | Not currently accessible |
|---|---|---|---|
| Standards bodies / multi-stakeholder initiatives (OECD.AI, ISO/IEC, IEEE, etc.) | | | □ |
| Specific civil society organizations | | | □ |
| Media / general public (people or communities indirectly impacted) | □ | | |
| Other | | | □ |

**Please describe how your organisation determines when and how reporting mechanisms are made accessible to different stakeholders (e.g. based on legal or regulatory requirements, internal policies, or risk level). Please also provide public URLs or details on how to submit reports. If none, write “Not applicable”**

Anyone can report a security vulnerability or AI-related risk through our public vulnerability disclosure policy (linked from https://suhul.us/.well-known/security.txt), which offers safe harbour for good-faith research and defines response times. Customers can also report issues through in-product support and support@suhul.us. Reports are triaged by the R&D lead by severity; confirmed issues are fixed, tested, and released through our normal deployment gate, and similar features are checked for the same issue. We do not currently share reports with third parties on a routine basis.

**4. If your organization has programs to incentivize risk(s) disclosure, incidents and systemic vulnerabilities, please share what type of programs you use?**

- Vulnerability Disclosure Programs (VDPs) with safe harbor / liability protection

**Please provide public URLs and/or details on your incentive program(s). If none, write “Not applicable.”**

Our public vulnerability disclosure policy (https://suhul.us/.well-known/security.txt) provides safe harbour for good-faith security research. We do not offer paid rewards.

**5. Does your organization use or contribute to the development of international technical standards, best practices or tools for the identification, assessment, and evaluation of risks?**

Uses

**If contributes or uses, please select all that apply:**

- NIST AI Risk Management Framework (AI RMF)
- Industry-led best practices or codes of conduct

**6. Does your organization collaborate with relevant stakeholders across sectors to assess and adopt risk mitigation measures to address risks that could have broad and potentially cascading socio-economic impacts? Please indicate which stakeholders you collaborate with, and whether**

**collaboration is on a standing basis (Always-on), activated based on context (Ad hoc/as relevant), or there is no current collaboration.**

| | Always-on | Ad hoc / as relevant | No current collaboration |
|---|---|---|---|
| Relevant employees (security, legal/compliance, etc.) | ☐ | | |
| Public authorities / regulators / AI safety institutes / other public sector bodies and inter-governmental coordination structures | | | ☐ |
| Industry partners or industry fora | | | ☐ |
| Users (end users, enterprise customers, developers using APIs) | | ☐ | |
| Business customers / deploying organizations (if different from “users”) | | ☐ | |
| Supply-chain partners (cloud, data providers, etc.) | | ☐ | |
| Academia / researchers | | | ☐ |
| Independent auditors | | | ☐ |
| Standards bodies / multi-stakeholder initiatives (OECD.AI; ISO/IEC; IEEE; AIID etc.) | | | ☐ |
| Independent civil society organizations (e.g., for evaluations, audits, or policy review) | | | ☐ |
| Media / general public (people or communities indirectly impacted) | | | ☐ |
| Other(s) (please add any other relevant stakeholder groups): | | | ☐ |
| Not applicable | | | ☐ |

> **Please describe examples of how your organization collaborates. Examples may include: participating in multi-stakeholder or industry initiatives; partnering with academia, civil society, or the public sector; establishing internal or joint governance structures with external experts or in partnership with other efforts to share knowledge and jointly address risks.**
>
> Collaboration is mainly with our own staff and customers: security, product and customer-facing staff review AI risks together; enterprise customers configure where their AI content may be processed and give feedback on AI outputs; and we work with our infrastructure and model providers under their security and no-training terms. We do not currently take part in multi-stakeholder AI safety initiatives; submitting this report is a first step.

**Any further comments and implementation documentation.**

Public references: AI modules at https://suhul.us/ai; privacy policy, model providers and the no-training commitment at https://suhul.us/en/privacy. We do not yet run a formal red-team programme or bias evaluation; both are planned within the next 12 months.

## Section 2 – Risk Management and Information Security

**7. Across the AI system lifecycle, what steps does your organization take to manage risks and vulnerabilities? Please indicate which specific practices, tools or metrics your organization uses at each lifecycle stage. You may refer to tools or metrics present in the OECD.AI Catalogue of Tools and Metrics or suggest additions. Respondents may complete only the lifecycle stages relevant to their organisation's role. Stages that do not apply may be left blank.**

> **Plan and design**

> **Collect and process data**

> **Build, integrate and interpret model**

> **Verify and validate**

> **Deploy**

> **Operate and monitor**

> **Please provide a consolidated description of your organisation's approach to risk management across the AI system lifecycle, including how the practices identified above operate together and any cross-cutting governance, review or escalation processes. Please also indicate how often these practices are reviewed and what types of events trigger reviews and updates. Where relevant, please distinguish between different types of AI systems (e.g. generative or agentic AI).**
>
> SUHUL is an application developer and deployer: we do not train models, so our risk management focuses on how third-party models are integrated into business workflows.
> **Plan and design:** each AI feature is scoped by what data it reads, what actions it can take and who can trigger it; features are grouped into read-only (summaries, classification), drafting (suggestions labelled as AI), and agentic (tool use). **Data:** customer content is routed by a single egress policy — by default it stays on Cloudflare Workers AI; other providers are used only after the customer opts in; Taiwan-only customers are processed only in the Taiwan region; identifiers are masked before leaving Cloudflare when the customer's policy requires it. **Build and integrate:** all model calls go through common entry points that apply the egress policy, usage metering and logging; untrusted content from tools and attachments is screened for prompt injection. **Verify and validate:** type checks and automated tests (including prompt-injection and permission cases) run in a clean build before every release. **Deploy:** releases are uploaded at 0% traffic and switched only after sign-off by the Governor. **Operate and monitor:** every AI call is logged with feature, model, workspace, operator and cost; outcomes and failures are recorded and feed alerts.
> For agentic features, each tool is registered with a permission and risk level, can be disabled centrally, and actions that change records or contact third parties require human confirmation bound to the exact

parameters; refunds, rate publication and financial close are blocked for AI agents. Practices are reviewed when a model or provider changes, when a new agentic capability is added, after any incident, and at least annually.

**Please describe any measures used to ensure data quality and fairness across the data collection and curation stage (e.g., dataset audits, filtering training data to remove harmful content or personal information, fine-tuning on curated data, use of moderation tools, assessment of limitations to model generalisability).**

We do not collect or curate training data and do not fine-tune models. Customer data is not used to train any model, and our model providers are contractually bound not to train on it. At inference time we apply data minimisation: identifiers (national ID numbers, email addresses, phone numbers, card numbers verified by checksum, and bank account numbers) are masked before content is sent to providers outside Cloudflare when the customer's policy requires it, and only the number of masked items is logged, never the original values. Content from tools, attachments and external sources is marked as untrusted. We have not yet carried out a formal fairness assessment of our AI features; this is planned within the next 12 months.

**8. What security measures are used in testing AI systems?**

- Locally hosted testing environments
- Containment measures for prompt-injection and jailbreak attempts
- Rate limits and resource caps
- Human-in-the-loop controls for higher-risk test cases
- Access controls, including roles-based permissions
- Audit logging and monitoring
- Use of synthetic, anonymized, or dummy data, as appropriate

**8.A. Please indicate, if relevant, the team or role responsible for overseeing secure testing and whether this responsibility is shared across multiple teams and/or functions.**

The R&D lead is responsible for secure testing. Tests run in isolated local and clean build environments with seeded dummy data, never against production customer data; production releases are uploaded at 0% traffic before sign-off.

**9. How does your organization protect intellectual property, including copyright-protected content?**

**9.A. Safeguards that promote lawful data use and transparency in AI training:**

- Not applicable

**9.B. Safeguards that reduce the risk of dissemination of outputs that may infringe copyright:**

- Designating clear points of contact so creators or rights holders can raise concerns about infringing outputs

**10. How do you protect privacy and prevent your systems from disclosing confidential or sensitive data across the AI lifecycle?**

- Make information about data privacy, practices, and controls accessible to a diverse set of stakeholders
- Use security controls, including encryption at rest and in transit
- Provide organizational customers with controls to configure and implement privacy choices (e.g., data retention, logging, access, or processing settings)

> **Describe key privacy measures and data-protection standards applied.**
>
> Customers choose where their AI content may be processed (global, or Taiwan region only) and whether it may be sent to providers outside Cloudflare; if the setting cannot be read, the strictest option applies. Identifiers are masked before external processing according to the customer's policy. Customer data is not used for model training. Every workspace is isolated by tenant identifier, and AI actions are checked against the caller's workspace before execution. Data is encrypted in transit and at rest on Cloudflare infrastructure. Our privacy policy (https://suhul.us/en/privacy) describes processing purposes, sub-processors and data subject rights, and we follow Taiwan's Personal Data Protection Act and applicable foreign requirements for our customers.

**11. How does your organization assess cybersecurity risks and implement policies to enhance the cybersecurity of advanced AI systems across the AI system lifecycle? Respondents may complete the checkbox subsections in 11.A–11.D or describe their overall approach under 11.E.**

> **11.A. Governance, policy and assessments**
>
> - Perform threat modeling that includes AI-specific risks (e.g., data poisoning, prompt injection)
> - Maintain or update security policies aligned with recognized standards (e.g., NIST CSF, ISO/IEC 27001, SOC 2, CIS Critical Security Controls), and review them regularly in response to emerging risks or regulatory guidance
> - Maintain vulnerability management processes (scanning, channels for external reporting, patching, remediation)
> - Implement incident response, disaster recovery, or business continuity plans, including escalation procedures for high-severity incidents
> - Integrate cybersecurity AI risks into broader cybersecurity risk management activities
> - Regularly review security measures and update them in response to emerging risks, incidents, or major releases
> - These responsibilities are delegated contractually to model or infrastructure providers

> **11.B. Technical and operational controls**
>
> - Encrypt AI assets (datasets, models, configurations) in storage and in transit as appropriate
> - Monitor system activity and log access to detect anomalies or misuse
> - Conduct secure AI development and/or deployment practices
> - These responsibilities are delegated contractually to model or infrastructure providers

> **11.C. Protection of sensitive or high value AI assets**

- Limit access to proprietary AI assets based on role or need-to-know
- Store model weights, datasets, and algorithms in secure, access-controlled environments
- These responsibilities are delegated contractually to model or infrastructure providers

**11.D. Insider risk protection detections**

- Monitor for unauthorized internal access or suspicious activity
- These responsibilities are delegated contractually to model or infrastructure providers

**11.E. Please describe the main security measures your organization applies and indicate which are specific to particularly sensitive AI assets, where relevant.**

Our services run on Cloudflare's serverless platform; physical security and infrastructure hardening are delegated to Cloudflare, and model hosting is delegated to the model providers under contract. On our side: tenant isolation on every query; role-based access control and audited headquarters access to customer workspaces; secrets stored in the platform secret store and never in code; a web application firewall in front of all systems; dependency vulnerability alerts; a single egress policy for AI content; prompt-injection screening and blocking of data exfiltration through external images and links; centralised usage limits for all AI calls; and a public vulnerability disclosure policy with security.txt. The most sensitive AI assets are model-provider API keys and customer AI content: keys are held only in the secret store, and customer content is subject to the egress and de-identification controls above.

**12. What is your organization's vulnerability management process? Does your organization take actions to address identified risks and vulnerabilities, including in collaboration with other stakeholders?**

- Internal vulnerability disclosure receipt and tracking processes
- Investigation of reported or identified risks and vulnerabilities
- Development of patches or other mitigations for confirmed risks or vulnerabilities
- Testing of patches or other mitigations
- Scanning for and mitigation of similar risks or vulnerabilities

**13. What post-deployment practices does your organization use to monitor and respond to vulnerabilities, incidents, emerging risks and misuse of advanced AI systems?**

- Conduct post-deployment monitoring
- Perform root-cause analysis and apply lessons learned to improve future development and/or deployment practices
- Monitoring and emergency controls specific to agentic AI systems (e.g., ability to interrupt autonomous action sequences, override tool-use permissions, or revoke delegated authority)

**14. Do you develop or deploy agentic AI systems (systems that can autonomously plan/act or use tools)?[4]**

Yes

**14.A. If yes, please indicate which categories of controls your organization applies. Please select all.**

- Constraints on the agent's action space (e.g., tool-use permissions, environment sandboxing, allowlists/blocklists of resources, scope limits per task)
- Human oversight or interruptibility mechanisms (approval requirements for sensitive actions, escalation paths, interruption interfaces)
- Logging of agent actions

**Please provide any further comments and implementation documentation.**

Public documentation: https://suhul.us/ai (AI modules), https://suhul.us/en/privacy (privacy policy, model providers, no-training commitment).

## Section 3 – Transparency Reporting on Advanced AI Systems

**15. Does your organization publish public reports and/or technical documentation related to the capabilities, limitations, and domains of appropriate and inappropriate use of advanced AI systems? Are these clear and understandable by non-technical audiences?**

Yes

**15.A. If yes, which of the following information is included in your organization's publicly available documentation for any of your AI products?**

*No answer provided*

**15.B. If yes, how often are such reports usually updated?**

*No answer provided*

**16. How does your organization share information with a diverse set of stakeholders (other organizations, governments, civil society and academia, etc.) regarding the outcome of evaluations of risks and impacts related to an advanced AI system?**

- Public documentation

**Describe how evaluation results are shared and with which stakeholders (e.g. other organizations, governments, civil society and academia, etc). You may also indicate whether non-technical or plain-language summaries are provided for broader audiences.**

We share risk-related information through public documentation written for non-technical readers: each AI module is described at https://suhul.us/ai with its purpose, inputs, outputs and pricing; our privacy policy explains which model providers are used, where AI content may be processed and that customer data is not used for training; and this HAIP report is published on OECD.AI. Enterprise customers receive security and AI questionnaires on request (including our CSA STAR self-assessment answers). We do not currently share formal evaluation results with governments or academia.

**17. Does your organization publish public privacy policies addressing the use of personal data, user prompts, and/or the outputs of advanced AI systems?**

Yes

> **If yes, provide links to privacy policies or notices covering personal data, user inputs, and AI-generated outputs.**
>
> Privacy policy (English): https://suhul.us/en/privacy — Section 12 covers AI processing: personal data, prompts and AI outputs are used only to provide the service, are not used to train any foundation or general-purpose model, and model providers are contractually bound not to train on them. Chinese version: https://suhul.us/legal.html#privacy

**18. How does your organization provide information about the sources of data used for the training of advanced AI systems, as appropriate, including information related to the sourcing of data annotation and enrichment?**

- Not applicable

> **Please provide links to any publicly available information.**
>
> Not applicable: SUHUL does not train models. Our AI modules: https://suhul.us/ai ; privacy policy: https://suhul.us/en/privacy

**19. Does your organization demonstrate transparency related to advanced AI systems through any other methods?**

Yes

> **If yes, describe additional method(s), including where relevant transparency measures related to AI agents**
>
> Each AI module is listed publicly at https://suhul.us/ai with its purpose, inputs, outputs and credit pricing, and there is a no-login trial so anyone can see how it behaves. Inside the product, AI suggestions and AI-generated assets are labelled, and AI outputs are recorded with their sources and classified as fact, inference, suggestion, pending approval or executed action. We also publish a security page and a vulnerability disclosure policy at https://suhul.us.

**20. Does your organization publish statistical insights derived from anonymised, aggregated usage data to advance understanding of AI adoption and its economic impacts in areas including occupations and tasks by region, sector and country? To this end, do you contribute to international cooperation efforts?**

No

**Please estimate the total staff time and annual expenses associated with conducting compliance activities and include any further comments and implementation documentation (optional).**

*No answer provided*

## Section 4 – Organizational Governance, Incident Management, and Transparency

**21. How has AI risk management been embedded in your organization's governance framework?**

- Executive oversight

> **Please describe how AI risk management is integrated into governance structures, including: which roles or bodies have advisory versus decision-making authority; which roles or bodies have final accountability for AI risk-related decisions; and how often policies are reviewed or updated (e.g., following major incidents, regulatory changes, or material updates to systems).**
>
> SUHUL is a small company. The Governor holds final accountability for AI risk decisions and signs off every production release before it receives traffic. The R&D lead is responsible for implementing AI safeguards (egress policy, de-identification, human confirmation of agent actions, logging) and advises the Governor on new AI features, model or provider changes, and incidents. AI risk management is part of the information security policy approved by the Governor in October 2026. Policies and safeguards are reviewed at least annually and additionally when a model or provider changes, when a new agentic capability is introduced, after an incident, or when relevant regulation or guidance changes.

**22. Are relevant staff trained on your organization's governance policies and risk management practices?**

No

**23. Are steps taken to address reported incidents documented and maintained internally?**

Yes

> **If yes, describe how incident records are kept and used for follow-up (e.g., incident logs, root-cause analysis reports, internal tracking systems). You may indicate whether and how you define escalation thresholds for when incidents must be surfaced to senior leadership or external parties.**
>
> Each incident and its fix is recorded internally with the cause, the change made, the tests added and the other features checked for the same pattern; the deployment history links each fix to its release. Lessons are written into internal engineering guidance so the same class of problem is checked in future work. A structured security incident register with severity, timelines and follow-up actions is being added to our administration console.

**24. Does your organization share research and best practices on addressing or managing risk?**

No

**25. Does your organization use international technical standards or best practices for AI risk management and governance policies?**

Yes

> **If yes, please indicate which standards or frameworks apply:**
>
> *No answer provided*

> **If yes, describe which international technical standards or best practices your organization uses.**
>
> NIST AI Risk Management Framework 1.0 and the Generative AI Profile (NIST AI 600-1), used for self-assessment; OECD AI Principles and the Hiroshima Process Code of Conduct; Cloud Security Alliance Cloud Controls Matrix and AI Controls Matrix, used for our STAR Level 1 self-assessments.

**Any further comments and implementation documentation.**

Formal AI governance training for staff is planned within the next 12 months; today, engineering staff follow written internal guidance on AI safeguards.

## Section 5 – Content Authentication & Provenance Mechanisms

**26. What mechanisms, if any, does your organization put in place to allow end-users, where possible and appropriate, to know when they are interacting with an advanced AI system developed by your organization?**

- Distinct branding or labeling for AI-generated content (e.g., “AI-generated” tag, metadata label)
- Documentation or FAQs explaining when and how AI is used in the system
- Disclosure to end-users when an AI system is acting autonomously or executing actions on their behalf (e.g., making purchases, sending communications, modifying files)

> **Provide any additional details or context on how or when users are informed they are interacting with advanced AI systems (e.g. if such mechanisms are planned, in testing, or applied only to certain products).**
>
> Our users are employees of business customers working inside SUHUL EIP. AI suggestions are shown with an "AI suggestion" label and an AI icon, AI-generated images in the asset library are marked with "AI-generated" as their source, and AI modules are documented at https://suhul.us/ai. When the AI assistant proposes an action on the user's behalf (for example creating a record or sending a message), it shows a confirmation card describing the exact action and parameters, and nothing is executed until the user approves. We are reviewing automated replies that reach our customers' own end customers so that they are consistently disclosed as AI-assisted.

**27. Does your organization use content provenance detection, labeling or watermarking mechanisms that enable users to identify content generated by advanced AI systems?**

Yes

**27.A. If yes, how?**

*No answer provided*

**27.B. If yes, does your organization use international technical standards or best practices when developing or implementing content provenance?**

*No answer provided*

**27.C. Does your organization collaborate on or invest in research to advance content authentication and provenance?**

*No answer provided*

**Describe how provenance mechanisms are implemented and which standards or tools are used, and any collaborations or investments to improve content authentication (e.g., C2PA credentials, joint R&D projects, open-source initiatives, standard-setting groups). You may also indicate whether these mechanisms are tested for robustness against removal, tampering, or circumvention.**

Provenance today is implemented at the application level: AI-generated assets carry an "AI-generated" source field in the asset library, AI outputs are logged with the model used and their sources, and business documents issued from SUHUL EIP carry a verification QR code that lets recipients check where the document came from. We do not yet embed watermarks or C2PA metadata in exported AI-generated images; this is planned.

**Any further comments and implementation documentation.**

SUHUL does not offer video or voice synthesis of real people.

## Section 6 – Research & Investment to Advance AI Safety & Mitigate Societal Risks

**28. Does your organization advance research and investment in the following areas?**

- Security
- Safety
- System performance and outcomes (e.g. accuracy, consistency)
- Robustness or reliability

**If yes, please describe key research or investment initiatives in the areas selected above (e.g., dedicated safety teams, funding of research programs, contributions to open-source safety tools, publishing research, partnerships and collaborative research initiatives). You may also indicate whether related outputs are publicly shared, shared with partners or within consortia or not publicly disclosed.**

As an SME we invest in applied engineering rather than foundational research. Security and safety: prompt-injection defences for agentic features, de-identification before model calls, in-country processing options and human confirmation for business actions. System performance: we evaluate speech-to-text and model outputs on our own Traditional Chinese business test sets before changing models or settings, and record AI outcomes in production to detect regressions. Robustness: when a provider fails, features fall back to an alternative provider where the customer's processing policy

> allows, and fail with a clear message where it does not. Socio-economic: our products are designed to make AI usable by small and medium-sized businesses in Taiwan and Asia, with a free plan and per-use credit pricing.

**29. Does your organization pursue research or investment related to socio-economic or environmental considerations in the development and deployment of AI?**

- Socio-economic considerations (e.g., AI adoption and productivity, workforce transitions, skills development, access to AI tools)

> **If yes, please provide examples of relevant research or investment initiatives for the areas selected above.**
>
> Our products are built to make AI usable by small and medium-sized businesses in Taiwan and Asia: SUHUL EIP has a free plan, AI modules are priced per use rather than by large subscriptions, interfaces are in Traditional Chinese and English, and customers can keep AI processing in Taiwan. Typical uses are document recognition, meeting minutes, customer replies and financial checks that reduce repetitive manual work for small teams.

**Any further comments and implementation documentation.**

Public documentation: https://suhul.us/ai

## Section 7 – Advancing Human and Global Interests

**30. Does your organization support any digital literacy, education or training initiatives to improve user awareness and/or help people understand the nature, capabilities and limitations of advanced AI systems?**

No

**31. Does your organization pursue AI projects that aim to deliver societal or public benefits, including through partnerships with civil society and community groups?**

No

**Any further comments and implementation documentation.**

We do not yet run formal AI literacy or public-benefit programmes. Our contribution today is making AI practical and affordable for small and medium-sized businesses in Taiwan and Asia — a free plan, per-use pricing, Traditional Chinese interfaces, and the option to keep processing in Taiwan — while keeping customers in control of where their data is processed and requiring a human decision for consequential actions.