

AI Governance Policy

How Workplacer develops, tests, operates, maintains, and retires the AI capabilities in its platform.

1. PURPOSE & SCOPE

This policy governs the AI systems used within the Workplacer platform across their lifecycle — creation, testing, operation, maintenance, and termination. Workplacer builds AI features on top of established third-party foundation models; it does not train its own foundation models. AI is used for document analysis and abstraction, insight and summary generation, semantic search over a customer's own content, and an in-app assistant.

2. PRINCIPLES

- **Human oversight** — AI outputs are advisory and support human decision-making; they do not take irreversible actions on customer data without user initiation or review.
- **Data minimisation** — only the data necessary for a given task is sent to a model.
- **No training on customer data** — customer personal data is not used to train external AI models.
- **Customer control** — customers can configure or disable AI processing of their data.

3. EU AI ACT CLASSIFICATION

Roles. Under Regulation (EU) 2024/1689 (the AI Act), Workplacer is the *provider* (Art. 3(3)) of the AI systems in its platform. Because those systems are built on general-purpose AI models from Anthropic (Claude) and Voyage AI (embeddings), Workplacer is also a *downstream provider* (Art. 3(68)). Our customers are the *deployers* (Art. 3(4)) of those systems within their organisations. Anthropic and Voyage AI have their own obligations as providers of general-purpose AI models (Chapter V). Those obligations are separate from the classification of our systems below.

Assessment. The AI Act sorts AI systems into four risk levels. Each Workplacer AI feature has been assessed against all four, and this section records the result (assessment dated 1 October 2026):

Risk level	Outcome for Workplacer
Prohibited (Art. 5)	No feature. We do no emotion recognition in the workplace, no biometric identification or categorisation, no social scoring, and use no manipulative or exploitative techniques.
High risk (Art. 6, Annexes I and III)	No feature. None is a safety component of a product covered by Annex I (Art. 6(1)), and none falls within an Annex III area (Art. 6(2)). In particular, no feature is used to recruit or select people, decide on terms of employment, promotion or termination, allocate tasks based on individual behaviour or personal traits, or monitor and evaluate the performance or behaviour of individuals (Annex III, point 4). No feature is a safety component in managing or operating critical infrastructure, such as heating or electricity supply (Annex III, point 2). Because no feature falls within Annex III, we do not rely on the Art. 6(3) exemption.

Risk level	Outcome for Workplacer
Limited risk (transparency obligations, Art. 50)	Two kinds of feature. The in-app assistant talks directly with users, so it is presented as an AI system (Art. 50(1)). Features that generate text, such as summaries, insights, narratives and report sections, label their output as AI-generated in the product and when it is shared outside it (Art. 50(2)).
Minimal risk (no specific obligations)	All other AI features: extracting data from a customer's own documents, classification and tagging, and semantic search. They follow the general controls in this policy.

Workplace survey data. Employee survey results — our own surveys and Leesman results delivered to a customer's workspace — are processed only in aggregate, by building, site and department. Groups of fewer than five respondents are never reported. No feature scores, ranks or profiles individual employees or infers emotions or personal traits, and survey data is never combined across customers.

Re-assessment. The classification is reviewed with this policy. It is also re-assessed before release whenever a new AI feature would process data about identifiable individuals (particularly employees) or would change the purpose of an existing feature. It is re-assessed as well whenever the Commission issues guidelines on high-risk classification (Art. 6(5)) or on transparency obligations (Art. 96), or approves a code of practice on marking AI-generated content (Art. 50(7)).

4. CREATION & DEVELOPMENT

- Models are selected against defined standards (higher-capability models for analysis, faster models for high-volume/classification tasks) and referenced by stable model aliases.
- System prompts are maintained centrally in the codebase and reviewed as code, not embedded ad hoc.
- Structured AI outputs are validated against defined schemas, with post-validation before results are used.

5. TESTING & EVALUATION

- AI features are covered by an automated evaluation suite, including scheduled (nightly) evaluations and an LLM-based judge tier, to detect quality regressions.
- Changes to AI features follow the standard development lifecycle — review and automated checks before deployment.

6. OPERATION & DATA HANDLING

Control	Description
Document data mode	Each customer selects how much document content the AI may access: full document, text only, or text with PII redacted. AI can also be disabled entirely.
PII scrubbing	Personal data is automatically scrubbed from text before it is sent for embedding/semantic search.
Review policy	An AI review policy setting governs whether AI-assisted checks are off, advisory, or blocking for the relevant workflow.

Control	Description
Usage metering	AI usage is metered per customer for transparency and control.
Rate limiting	AI endpoints are rate-limited to prevent abuse and runaway cost.
Tenant isolation	AI operates only on the requesting company's data; retrieval is scoped to that company.

7. PROVIDERS & TRANSFERS

AI processing is provided by Anthropic (Claude models) and Voyage AI (text embeddings). These providers are US-based; transfers of personal data are governed by EU Standard Contractual Clauses. Personal data is not used by these providers to train their models.

8. TRANSPARENCY & AI LITERACY

- **Transparency to people (Art. 50(1) and (2))** — AI is identified by one consistent mark (◆) wherever it is used in the product. Controls that start AI work carry the mark. AI-written text carries an "AI-generated · Review before relying on it" line. Values that AI read from a document or pre-filled carry the mark until a person confirms or edits them. The in-app assistant is labelled "AI assistant" and shows a standing notice that it can make mistakes and that data should be validated before it is relied on. Each mark links to a plain-language explanation of what the AI does, that it can be wrong, and how to switch it off. Shared links and exported reports carry the same visible labelling.
- **Machine-readable marking (Art. 50(2))** — AI output is also marked in a form software can detect. We use the IPTC Digital Source Type vocabulary, the one used by C2PA content credentials and major platforms. `trainedAlgorithmicMedia` means generated by AI. `compositeWithTrainedAlgorithmicMedia` means AI output edited by a person, or content that mixes AI and human input. The marking is applied in every channel AI output leaves through:
 - `schema.org` markup and `data-ai-*` attributes on AI content in the product;
 - `<meta>` tags, JSON-LD and `AI-Generated` response headers on shared links;
 - XMP and document metadata in exported report PDFs and presentations;
 - `X-AI-*` headers on emails that carry AI text;
 - a `provenance` field on AI items in the public API.

An automated check in our build pipeline fails if a channel loses its marking. Plain-text watermarking is not yet technically robust, so metadata-based marking is applied "as far as technically feasible". We will align with the Commission's code of practice on marking AI-generated content (Art. 50(7)) once it is published.

- **Documentation** — the AI providers, data handling and customer controls are documented in this policy, our Data Processing Agreement and our public Trust page.
- **AI literacy (Art. 4)** — everyone at Workplacer who builds, operates or supports AI features completes our internal AI literacy guidance: what the models do and cannot do, how outputs can be wrong, how outputs are reviewed, and how customer data is protected. Completion is recorded, and the guidance is refreshed when this policy changes and at least annually.

- **Customer enablement** — product guidance explains what each AI feature does and that its output is advisory, so deployers can meet their own AI literacy and oversight obligations.

9. MAINTENANCE

Models are referenced by stable aliases so provider updates are adopted in a controlled way; AI feature quality is monitored via the evaluation suite and production error monitoring. Prompts and model choices are revised as features evolve, under the standard review process.

10. TERMINATION

AI processing can be disabled by a customer at any time via the platform's AI settings. On account termination, customer data is returned or deleted per our Data Processing Agreement; no customer data is retained by AI providers for training.

11. REVIEW

This policy is reviewed at least annually and when AI capabilities, providers, or applicable regulation change materially.

This policy should be read alongside the Information Security Risk Management summary, Third-Party (Subprocessor) Management Policy, Data Processing Agreement, and Privacy Policy.