AI Compliance FAQ

This document is intended to answer frequently asked questions about Culture Amp's responsible AI practices. The answers are provided for guidance purposes only and do not form part of any contractual agreement. While we endeavour to keep our FAQs up to date, the information is accurate as of the date listed in the footer. If you have any questions, please reach out to your Customer Success Coach or Account Manager.

Model Training & Data Usage

Is customer data (inputs, prompts, feedback) used to train AI models?

As a rule, Culture Amp never uses *identifiable* customer data to train any AI models. If we ever choose to train our own models we would exclusively use *de-identified* customer data only. To add some additional color:

Third-party Proprietary LLMs

Some of Culture Amp's AI-powered features rely on third-party proprietary LLMs – such as Gemini from Google and Claude from Anthropic. Culture Amp does not allow any providers to use customer data to improve their proprietary models. Specifically:

- ☐ Culture Amp has contractual agreements ensuring that third-party providers cannot train their underlying models on customer data.
- ☐ Customer data is processed ephemerally, solely to generate the outputs requested at the time of use, and is not retained by the third-party provider.

Developing Culture Amp's AI models

We are exploring a future state where Culture Amp may develop its own AI models to better serve customer needs. Any such work would fall under our license for Service Improvements, as outlined in Clause 5.4 of the General Terms. Under this clause, Culture Amp may develop and improve models using de-identified and/or aggregated customer data. This work:

- ☐ Would only use data that cannot directly or indirectly identify a customer or user
- ☐ Would not enable reconstruction of individual responses or customer-specific information

Can customers opt-out of their de-identified or aggregated data being used to train Culture Amp's AI models?

No. While Culture Amp is not currently developing its own AI models, if we were to do so in the future, using de-identified customer data would be essential to ensure the accuracy, reliability, quality, and trust of the AI product experience for all customers. At this time, we do not offer an enterprise option to exclude your data from this de-identified pool.

How do you de-identify customer data before it is used?

De-identification measures may include:

- □ We classify and detect direct identifiers - such as names, emails, and dates of birth - using regular expressions, conditional logic, and data profiling. Once identified, these values are secured through a combination of redaction (removal) or cryptographic hashing (SHA-256) so that original values cannot be reconstructed.
- □ Conditional selection of demographic, and organizational data, where combinations of attributes could increase identification risk.
- □ Use of paired or derived columns, where original values are replaced with non-identifying representations suitable for analysis.
- □ Aggregation and segmentation, so that analysis focuses on patterns across groups rather than individual records.

Data Privacy & PII

Is Personally Identifiable Information (PII) redacted or masked before being sent to third-party proprietary AI models?

No. Text is passed as written to AI models so that features like Coach can provide personalized responses. If you are concerned about your employees inputting sensitive information into our AI features, such as Coach, we recommend setting internal guidelines that clarify what information may be shared. Administrators also have the option to disable certain AI features, which prevents employees from accessing the AI and therefore from entering any information into the tool.

Does Culture Amp allow proprietary third-party model providers (e.g., Google, Anthropic) to retain customer data?

No. Customer data processed by third-party AI providers is handled ephemerally—meaning it is only used to generate outputs requested at the time of use and is not retained by the provider. Additionally, Culture Amp has contractual agreements in place to ensure that third-party providers cannot use customer data to train their underlying models.

Can customers prevent certain personal data (PII) fields from AI processing?

No. Culture Amp does not currently offer customer-configurable PII filters. However, our AI features are designed to balance personalization with data minimization. For example, the Comment Summaries feature does not use any structured identifiers such as names, email addresses, or employee IDs. In the interest of minimizing data exposure, only the free-text question and content are processed.

Data Retention

How long are user prompts and AI-generated responses retained?

Culture Amp stores Customer Data as long as necessary to provide our services or until a formal deletion request is received. Customers may exercise their right to erasure at any time by emailing privacy@cultureamp.com.

Third-Party Processors

What subprocessors are used to support your AI-features?

Culture Amp maintains a comprehensive list of authorized subprocessors, which is available for your review lat https://www.cultureamp.com/company/legal/subprocessor-list.

Where Culture Amp's AI-features rely on third-party proprietary models – where are those models hosted, and where is customer data processed?

When Culture Amp's AI features rely on third-party proprietary LLM models, those models may be hosted in locations as documented in the subprocessor list made available on Culture Amp's website. For clarity, this may differ from your storage location. That said, in all cases there is a secure communication mechanism and ephemeral use which means the third-party model will not store Customer Data or use it for training purposes.

Nevertheless, we aim to prioritise processing data in your chosen data residency region (Australia, the EU, or the US). However, in some cases data may be transmitted internationally for transient processing. For example, to maintain service continuity or to access advanced model capabilities not yet available in a specific local region.

All such activity is governed by our Data Processing Addendum (DPA), ensuring consistent global protection and privacy safeguards.

Security & Infrastructure

Are the AI models susceptible to prompt injection or adversarial attacks?

We use a layered defense-in-depth approach against AI-specific threats like prompt injection, combining guarded prompts via an AI gateway, model/vendor safeguards, input/output filtering, AI guardrails, strict data access controls, monitoring/detection, and adversarial testing aligned to OWASP Top 10 for LLMs and Agentic applications. This is in addition to the robust SLDC controls already in place.

Has a penetration test specifically for AI features been conducted?

Culture Amp conducts annual third-party penetration testing across its entire platform. Additionally, new features, including AI features, undergo separate one-off penetration tests prior to initial release. Summarized attestations of independent testing is available at security.cultureamp.com.

AI Governance

How does Culture Amp ensure Responsible AI practices are embedded at the organisation?

Culture Amp is committed to the responsible development and use of AI, guided by a robust framework of compliance, governance, and ethical principles. Our approach includes:

- ☐ **Our Responsible AI (RAI) Framework**, which informs the ethical design, development, and deployment of AI across our platform. It aligns with regulations such as the EU AI Act and is grounded in principles of accountability, fairness, transparency, and reliability.

- ☐ **An AI Governance Committee**, operating within our Risk Management Committee, which oversees the development and use of AI systems. This committee is responsible for setting and maintaining AI-related policies and standards, assessing risks and impacts, guiding AI initiatives, and ensuring legal and regulatory compliance.

- ☐ **A cross-functional working group**, comprising experts from across the business, that supports Culture Amp's AI efforts by advising on risk, championing safety-by-design, and helping to embed security and ethical safeguards.

To affirm Culture Amp's approach was consistent with industry standards, Culture Amp sought and obtained ISO 42001 compliance.

What does it mean to be ISO 42001 compliant?

ISO/IEC 42001:2023 is the first international standard for AI management systems (AIMS), providing a framework for organizations to develop, deploy, and use AI responsibly. It focuses on risk management, ethical guidelines, and transparency to ensure AI systems are safe, trustworthy, and compliant. To obtain the certification, Culture Amp had to prove to an independent third-party that responsible AI was embedded in our governance practices.

Legal

Who owns the Intellectual Property (IP) of the AI outputs?

While customers do not own Culture Amp's intellectual property—such as our AI algorithms or any modifications derived from them—they do retain full ownership of any Customer Data contained within AI-generated outputs. This is outlined in Clause 5.2 of our General Terms.

Does the solution comply with the EU AI Act?

Culture Amp is committed to regulatory compliance across all jurisdictions where we operate. Under Clause 9.1(b) of our General Terms, we warrant that our platform complies with all applicable laws and regulations, which includes the EU AI Act.

Is Culture Amp considered high-risk under the EU AI Act?

To the best of our knowledge, Culture Amp's AI features are not classified as high-risk under the EU AI Act. Culture Amp arrived at this conclusion after reviewing guidance from the European Parliament, relevant standards bodies, and industry peers. We have also sought advice from external legal counsel. Based on this review – and external opinion – we have confirmed that, at present, Culture Amp does not fall within the scope of a high-risk classification.

This assessment is based on the fact that our AI-related systems are not intended to be used for any of the purposes set out in Article III of the EU AI Act, including making decisions that affect employment terms—such as promotion, termination, or task allocation—or for monitoring or evaluating employee performance or behaviour.

That said, we expect our customers to exercise appropriate human oversight and judgment when using our AI features and to ensure their use remains compliant with all applicable laws. We will continue to monitor regulatory developments and update our position as guidance evolves.

Does Culture Amp have an AI Addendum to its General Terms?

No. We do not require a separate AI Addendum because the protections governing the use of AI are already fully integrated into our contract suite – including our General Terms, Privacy Policy and Data Processing Addendum (DPA).

Bias & Explainability

How do you test for and mitigate bias in Culture Amp's AI Coach product?

Culture Amp has and continues to test for harmful biases with our AI Coach product. AI Coach is built with a multi-layered approach to protect against and evaluate for bias including content guardrails, model selection, carefully engineered prompts, evaluations from human people scientists and automated LLM-as-judge testing, as well as the provision of user feedback mechanisms.

Our bias evaluation approach combines qualitative and quantitative evaluation methods.

- □ **Scenario-Based Testing:** Multi-turn conversation scenarios that stress-test the system's ability to avoid reinforcing biased assumptions
- □ **LLM-as-a-Judge Evaluation:** Automated assessment against defined bias criteria
- □ **People Science Expert Review:** Human testing and evaluation by our People Scientists

We evaluate whether AI Coach avoids demographic assumptions, excludes discriminatory language, avoids discriminatory recommendations, appropriately redirects biased inputs, and promotes inclusive approaches.

Are outputs from your AI features explainable?

Culture Amp is committed to ensuring that our AI-powered features are transparent, purpose-specific, and understandable for end users.

Our user experience design clearly indicates where AI is being used—for example, through the AI sparkle icon and in-product disclosures—so users know when content has been AI-generated and how it supports their workflow.

Where appropriate and technically feasible, we also provide visual traceability, such as showing the original content that contributed to an AI-generated summary. This helps users better understand how outputs are produced.

Internally, we maintain observability logging, which captures structured data about how AI features behave. This allows us to monitor, review, and trace AI system behavior over time.

How do you mitigate hallucinations in your AI features?

To help to prevent hallucinations, Culture Amp applies a multi-layered approach focused on AI feature design, contextual data validation, user control, testing, evaluation scoring of responses, third party model safeguards, and feedback mechanisms.

Users are always in charge: AI-generated content is clearly marked and presented as suggestions, and not decisions. The interface is designed for transparency, and users are prompted to review and take responsibility for the final output.

By clearly focusing the AI feature, grounding responses in trusted source material, and applying controls we reduce the likelihood of the model generating unsupported or inaccurate information. These measures are combined with our ongoing governance program to assess risks and benefits.

Controls & Functionality

Is there a "Human-in-the-Loop" mechanism?

Yes. Our AI features are designed to support - not replace - human judgment. At no point does AI replace human decision-making. Your employees remain in the drivers seat and are empowered to interpret, apply, or disregard AI-generated suggestions based on their own content and expertise.

Can we opt-out of Culture Amp's AI features?

No. AI is an integral part of the Culture Amp platform. It powers features that enhance the user experience and also supports core functionality that delivers the insights and reporting our platform provides. While customers can control how they interact with some AI features, certain foundational analytics functions are essential to the platform and cannot be turned off.

More specifically, analytical AI, such as Topic and Sentiment classification, is a core part of our reporting engine. By analyzing employee feedback in the context of both survey questions and responses, it improves accuracy and correctly interprets mixed or neutral sentiment. Because these analytics are essential to generating meaningful insights, they cannot be disabled.

However, administrators and users have some control over generative AI features, such as AI Coach. Administrators can enable or disable features like AI Coach and Comment Summaries, which prevents employees from accessing these functionalities if disabled. Users can also choose not to interact with AI Coach by simply not entering any prompts into the assistant.