

This document describes how Scorafy uses artificial intelligence to assess written, audio and video responses, what controls sit around that use, where data goes, and what Scorafy does not claim. It is prepared for customer legal, procurement and data-protection review and is consistent with the published Scorafy Terms of Service, Data Processing Agreement, Privacy Policy and Security and Trust pages at scorafy.com. Where this document and the DPA differ, the DPA governs.

Provider	Cognitiv Pty Ltd, trading as Scorafy
ABN	51 688 133 977
Contact	adam@scorafy.com
Applies to	scorafy.com and au.scorafy.com
Review	Annually, or on any material change

01 WHAT THE AI DOES, AND DOES NOT, DECIDE

Scorafy uses a large language model to read each respondent's answers against the rubric, scoring rules and methodology context the customer has defined, and to produce a structured, evidence-cited report. The model assesses against the customer's criteria. It does not set the criteria, and it does not make the final decision.

STEP 1

Customer defines

Rubric criteria, level descriptors, scoring rules and context are set by the customer's administrators.

STEP 2

Respondent answers

Each person answers in their own words. Respondents are told before they start that AI is used, and can request a human alternative.

STEP 3

AI assesses

The model reads the complete answers with the rubric and context, scores each criterion and cites the answer text as evidence.

STEP 4

Human releases

Nothing is released to a respondent until an assessor releases it. Any score can be overridden.

02 HUMAN OVERSIGHT

Review before release

Respondents see no score, report or PDF until an assessor releases the result. The release records who released it and when, and is written to an audit log. Building an assessment and releasing its results are separable across roles, so the same person need not do both.

Confidence and escalation

Every AI evaluation records a per-criterion confidence. Low-confidence results, near-pass results and assessor disagreements are flagged for human review rather than passed through.

Overrides preserved, not overwritten

An assessor can override any AI score. The original AI value is stored beside the human-final value, so the audit trail shows both. The AI report itself is never mutated by an override.

Exportable evidence chain

Any assessment can be exported as an audit pack: a PDF and JSON pair carrying the model and configuration versions in force at evaluation time, the full override ledger and the release chain, readable without a Scorafy account.

03 THE MODEL AND HOW IT IS USED

Model

Reports are generated by Anthropic's Claude (currently Claude Sonnet 4.5) through Anthropic's commercial API. The model version is pinned and recorded on every report. Any change of model version, AI provider or inference region is notified to customers at least 30 days in advance.

Prompt-injection hardening

Everything a respondent submits, including uploaded document text and transcripts, is wrapped in escaped sentinel markers before it reaches the model, so instructions hidden inside submitted content are treated as data, not commands.

No training on customer data

Anthropic API inputs and outputs are deleted within 30 days (longer only if flagged for a usage-policy review, up to 2 years, or if required by law) and never used for training. Every transcription request is opted out of Deepgram's model-improvement programme.

Withheld scores stay withheld

Where an assessment withholds a score pending sign-off, the score is withheld server-side and kept out of the AI's written prose, so it cannot be recovered from the network response or inferred from the wording.

04 WHERE DATA GOES: SUB-PROCESSORS

Scorafy runs two isolated production environments: Dublin, Ireland for scorafy.com and Sydney, Australia for au.scorafy.com. Storage and application compute for a customer stay in the environment that customer uses. AI inference, transcription and email are processed in the United States, in transit. The full list is published in the Scorafy Sub-processor Register.

SUB-PROCESSOR	REGION	PURPOSE
Supabase	Dublin, Ireland or Sydney	Database, file storage and authentication. Primary data store.
Anthropic (Claude)	United States	AI report generation. Not used for training; deleted within 30 days.
Deepgram	United States	Transcription, only where audio or video is enabled. Opted out of model improvement.
Vercel	Dublin or Sydney (functions); global edge	Application hosting; no data stored.
Resend	United States	Transactional email.
Stripe	United States / Ireland / Australia	Subscription billing, where card billing is used.
GitHub	United States	Code hosting; encrypted backups (30 days).

Web analytics runs on marketing pages only, after cookie consent, and is never invoked for assessment data, responses or report contents.

05 RETENTION, DELETION AND SECURITY

Configurable retention, enforced by the system

Respondent personal data is retained for a period the customer sets per assessment, from 30 days to a year, enforced by scheduled jobs and logged. Media can be deleted at submission, and an organisation can switch audio and video off entirely. At the end of the agreement, data is deleted or returned at the customer's choice.

Tenant isolation and access control

Every table carries row-level security scoped to the organisation at the database layer. Five roles gate what each team member can do. Multi-factor authentication is available on every account, enforced server-side once enrolled, and can be required for every member of an organisation. Platform administration is a separate, fail-closed, audited role.

05 RETENTION, DELETION AND SECURITY (CONTINUED)**Encryption and breach notification**

Data is encrypted in transit (TLS) and at rest by the infrastructure sub-processors. Scorafy will notify the customer without undue delay, and in any event within 72 hours, after becoming aware of a personal data breach affecting customer personal data.

Respondent transparency

Respondents are told before they start that AI is used to analyse their responses. Owners can publish a custom AI notice on the landing page, and respondents can request a human alternative without first starting the assessment.

06 EU AI ACT POSITION

Assessment-style AI used to evaluate employees or learners is treated as a high-risk use under the EU AI Act, with obligations applying on the dates set by the Act as amended. Scorafy is built so the operating organisation can demonstrate meaningful human oversight on every individual outcome, not only at policy level: review before release, preserved overrides, per-criterion confidence, respondent notice and an exportable evidence chain. Conformity assessment and any external certification are activities for the deploying organisation; Scorafy supplies the technical and procedural building blocks and the documentation above.

07 WHAT SCORAFY DOES NOT CLAIM

- Scorafy does not hold SOC 2 or ISO 27001 certification and has not undergone an equivalent audit at its current stage. The controls above are real; the certificates are not yet.
- Scorafy is not "GDPR certified". GDPR is not a certification. The GDPR page describes the actual controls and every sub-processor.
- Single sign-on (SAML) is available on request; it is not part of the standard product today.
- Scorafy does not detect AI-written answers and does not claim to. Human review, evidence citation and assessor flags are the controls offered; any future integrity feature will be documented here before customers are asked to rely on it.
- AI cannot read context the way a human can, language and cultural bias exist in any model, and subjective assessment is inherently difficult. Scorafy is a structured first draft for a human assessor, not a final word.

Questions or a vendor questionnaire? Send it to adam@scorafy.com. Answers are drawn from the same published sources as this document, and anything not listed here is answered as "not held at this stage" rather than inferred.

08 DOCUMENT CONTROL

ITEM	DETAIL
Document	Scorafy AI Use Policy, SCF-POL-AI-2026-09, version 1.1 (supersedes version 1.0 of 8 September 2026)
Issued	27 September 2026
Owner	Adam Broons, Founder, Scorafy (Cognitiv Pty Ltd)
Sources	scorafy.com/trust · /security · /privacy · /terms · /dpa · /gdpr , and the Scorafy Sub-processor Register, as at the issue date
Next review	September 2027, or earlier on any change of model, sub-processor, region or retention behaviour
Precedence	Where this document and the Data Processing Agreement differ, the DPA governs