

NIST AI RMF Self-Assessment: cascadiaen

Self-assessed against NIST AI RMF 1.0 (72 subcategories). ISO 42001 cross-referenced via NIST's published crosswalk (NIST-to-ISO 42001, NIST AIRC). This is a transparency document and not a claim of certification.

TRIAGE CATEGORY KEY

0: N/A	Categorically inapplicable, not a function of build type or scale. Expected to be rare.
1: Not relevant for this build type	Applies to AI systems generally, not to this kind of build, regardless of scale.
2: Not relevant at this scale	Presupposes organisational structures not present in this operating model. May become applicable under a commercial build; revisit if the operating model changes.
3: Relevant and inherently met	Applies, and already satisfied by how the build is structured.
4: Relevant and required	Applies and was not already satisfied by how the build is structured. Each one was identified as an open item and explicitly addressed; the Addressed column records the outcome for each.

Function	Total	Cat 0 N/A	Cat 1 Not Relevant (Build Type)	Cat 2 Not Relevant (Scale)	Cat 3 Inherently Met	Cat 4 Required	Relevant subcategories met	Required items addressed
GOVERN	19	0	0	4	12	3	Complete	Complete
MAP	18	0	0	3	12	3	Complete	Complete
MEASURE	22	0	1	1	17	3	Complete	Complete
MANAGE	13	0	0	0	11	2	Complete	Complete
TOTAL	72	0	1	8	52	11	Complete	Complete

v1.1 | September 2026 | cascadiaen | Framework: NIST AI RMF 1.0 + ISO 42001 crosswalk (NIST-published) | GenAI risks: NIST AI 600-1 (July 2024) | v1.0 June 2026

Note: NIST AI 600-1 §3 (July 2024) offers optional suggested actions for a subset of subcategories, specific to generative AI. These are supplementary reference material, not self-assessment criteria; the triage categories above are the actual assessment. See Assessment col 1 (GAI Risk Tags) for a pointer to which subcategories it addresses: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf>

Terms used in this assessment

Run harness	The code that wraps every model and tool call: it checks schema, provenance, permissions and budget, and can halt a run.
HITL gate	A human-in-the-loop gate. A point the workflow cannot pass without a person acting.
Halt and flag	A halt stops the run. A flag records a concern on the output and lets the run continue.
Evals	Repeatable tests of model output against a fixed set of cases, held as a golden set by the human reviewer.
Decision record	A numbered, dated record of a decision taken about the build, held with the specification.
Triage category	How each NIST subcategory was classified for this build. See the key above.

NIST AI RMF Self-Assessment: cascadiaen												
Function	Category ID	Category Description	Sub ID	Subcategory Text (NIST AI 100-1)	ISO 42001 Clause Refs	ISO 42001 Annex A Controls (B.x.x)	ISO 42001 Annex A Category	GAI Risk Tags (NIST AI 600-1 §3) See Summary for ref.	Triage category	Notes	Evidence reference	Addressed
GOVERN	GOVERN 1	Policies, processes, procedures, and practices across the organization related to the mapping, measuring, and managing of AI risks are in place, transparent, and implemented effectively.	GV-1.1	Legal and regulatory requirements involving AI are understood, managed, and documented.	4.1; 6.2	B.2.2; B.2.4	A.2	DataPriv; HarmBias; IP	4 - Relevant and required	Applies (v1.0 triage 2026-07-02, superseded below): public deployment, third-party API terms, data licences; no personal data processed. Not yet consolidated - requires a light legal/licence note (API ToS, per-source data licences, no-personal-data statement), Pairs with GV-1.2 policy artefact. Addressed 2026-07-02: programme-level AI Policy & Legal Note authored - API terms (paid tiers), per-source data licences, no-personal-data statement, output disclaimers, AI-regulation position, IP/licence position (note s3). Reopened and re-addressed 2026-09-16 (v1.1); the note's s3 revised for client engagements - client-supplied material and personal data (now governed by the Client Data Handling and Privacy Policy), the nature of outputs with a standard limitations statement on the briefing note, AI-regulation positioning for a commercial service in the UK, EU and North America, data licences including web sources drafted per case, and ownership of client deliverables.	Partial: spec s5.2 (per-source licence/usage field), s21 (deployment). Gap: consolidated statement. ADDRESSED 2026-07-02: the AI policy and legal note (programme-level). RE-ADDRESSED 2026-09-16: the AI policy and legal note v1.1 s3; the Client Data Handling and Privacy Policy v1.0 (s2, s9, s10); the limitations statement on the briefing note, the run's caveat also in the QA report (implemented in system revision 16, item 29).	Yes
	GOVERN 1		GV-1.2	The characteristics of trustworthy AI are integrated into organizational policies, processes, procedures, and practices.	4.4; 5.2	B.2.2; B.2.3; B.6.1.2; B.6.1.3; B.9.3; B.10.3	A.2; A.6; A.9; A.10	CBRN; DataPriv; InfoInteg; InfoSec; IP	4 - Relevant and required	Substantively integrated in structure (policy-as-artefact layer, transparency, QA) but the explicit AI-policy statement (ISO A.2 - flagged in Ref Guide as category-4 candidate) does not yet exist. One short published policy note closes it; combine with GV-1.1. Addressed 2026-07-02: explicit AI policy authored - the seven NIST trustworthy-AI characteristics each mapped to a commitment and where it is enforced in the engine; governance/accountability stated (note s2). Reviewed 2026-09-16 (v1.1): s2 corrected to the system as built - access through private links, overrides recorded with the signatory's justification, explainability through the evidence cards, inputs register and briefing note, client data under the companion policy. The engagement folder, hold period and deletion are system revision 16, item 28.	Structure: spec s1.5, s10.3, s13. Gap: explicit AI policy artefact. ADDRESSED 2026-07-02: the AI policy and legal note (programme-level). REVIEWED 2026-09-16: the AI policy and legal note v1.1 s2. Engagement folder, hold period and deletion: system revision 16, item 28.	Yes
GOVERN	GOVERN 1		GV-1.3	Processes, procedures, and practices are in place to determine the needed level of risk management activities based on the organization's risk tolerance.	6.1.1; 6.1.2; 6.1.3			CBRN; Confab; Dangerous; HarmBias; InfoInteg; InfoSec; Obscene; ValueChain	4 - Relevant and required	Risk-management depth is implicitly calibrated (halt vs flag, core-friction-first evals) but risk tolerance itself is not documented. Requires a short risk-tolerance statement; pairs with MP-1.5 (one artefact covers both). Addressed 2026-07-02: risk-tolerance statement authored (note s4); its calibration paragraph ties the level of risk-management activity to the stated tolerances - enforcement heaviest where tolerance is lowest. Reviewed 2026-09-16 (v1.1): stances unchanged; s4 premises and enforcing mechanisms restated as built for client engagements (private links, commercially sensitive client figures held under the companion policy, spend ceilings per run and per case authoring). The engagement folder, hold period and deletion are system revision 16, item 28.	Partial: spec s11.4 (halt/flag), s19. Gap: documented risk tolerance. ADDRESSED 2026-07-02: the AI policy and legal note s4 (risk tolerance). REVIEWED 2026-09-16: the AI policy and legal note v1.1 s4; the Client Data Handling and Privacy Policy v1.0. Engagement folder, hold period and deletion: system revision 16, item 28.	Yes
GOVERN	GOVERN 1		GV-1.4	The risk management process and its outcomes are established through transparent policies, procedures, and other controls based on organizational risk priorities.	6.1.2; 6.1.3; 8.3			CBRN; Dangerous; DataPriv; HarmBias; Obscene; CivilRights	3 - Relevant and inherently met	This self-assessment + the spec's enforcement design are the transparent risk-management controls: halt/flag criteria, harness checks, QA always visible.	This workbook; spec s11.4, s12, s18.1 (QA visible).	N/A
GOVERN	GOVERN 1		GV-1.5	Ongoing monitoring and periodic review of the risk management process and its outcomes are planned and organizational roles and responsibilities clearly defined, including determining the frequency of periodic review.	8.2; 8.3; 8.4			Human-AI; InfoInteg; InfoSec; IP	3 - Relevant and inherently met	Review cadence documented; at each material change to the system or its operating model (for example a new capability, a change of hosting or AI provider, or the first client engagement) and annually (Lifecycle Map stage 6; policy note header). The first review against the built system ran 2026-09-16 (v1.1); none ran at the 2026-07-31 deployment. Roles collapse to the builder at this operating scale.	Lifecycle Map sheet (stage 6); the AI policy and legal note v1.1 header and s2 (review cadence); working agreement.	N/A
GOVERN	GOVERN 1		GV-1.6	Mechanisms are in place to inventory AI systems and are resourced according to organizational risk priorities.		B.4.2; B.4.3; B.4.4; B.4.5; B.4.6	A.4	DataPriv; Human-AI; InfoInteg; InfoSec; IP; ValueChain	3 - Relevant and inherently met	System/component inventory exists: the as-built dependency inventory, source registry and skills registry; the build programme documents inventory the builds themselves.	The as-built dependency inventory; spec s2 (the planned stack, June 2026); s5.2 (source registry), s8.2 (skills registry).	N/A
GOVERN	GOVERN 1		GV-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.		B.6.2.6	A.6	Human-AI; InfoSec; ValueChain	3 - Relevant and inherently met	Decommissioning is low-risk by design: no user accounts; access links can be killed individually and the service disabled; deleting the service and its data volume removes every run record made on the deployment (the retention of client material while operating is the MS-2.10 item); API funding caps bound any runaway.	The AI policy and legal note v1.1 s6 (killable links); spec s21 (decision record 16) for the funding cap; s17.3 (no multi-tenancy); MG-2.4; MS-2.10. The Client Data Handling and Privacy Policy v1.0 s8 (deletion; system revision 16, item 28).	N/A
GOVERN	GOVERN 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.	GV-2.1	Roles and responsibilities and lines of communication related to mapping, measuring, and managing AI risks are documented and are clear to individuals and teams throughout the organization.	5.3; 7.1; 7.2; 7.3; 7.4; 9.1	B.3.2	A.3	CBRN; Dangerous; HarmBias; Human-AI; InfoSec; ValueChain	3 - Relevant and inherently met	Solo operating model; roles and decision rights explicitly documented via the decision protocol ([Autonomous]/[Consult]/[Material]) and working agreement.	working agreement; spec s0 (decision classification).	N/A
GOVERN	GOVERN 2		GV-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.	7.2			Not covered by NIST AI 600-1 §3	2 - Not relevant at this scale	Presupposes personnel/partners to train. Solo builder; revisit if the operating model gains staff or partners.	-	N/A
GOVERN	GOVERN 2		GV-2.3	Executive leadership of the organization takes responsibility for decisions about risks associated with AI system development and deployment.	5.1; 5.2; 9.3.1; 9.3.2; 9.3.3			Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	The builder is the accountable decision-maker; material risk/deployment decisions are reserved to the human by protocol and cannot be resolved silently.	working agreement decision protocol; spec s0 [Material] register.	N/A

NIST AI RMF Self-Assessment: cascadiaen												
Function	Category ID	Category Description	Sub ID	Subcategory Text (NIST AI 100-1)	ISO 42001 Clause Refs	ISO 42001 Annex A Controls (B.x.x)	ISO 42001 Annex A Category	GAI Risk Tags (NIST AI 600-1 §3) See Summary for ref.	Triage category	Notes	Evidence reference	Addressed
GOVERN	GOVERN 3	Workforce diversity, equity, inclusion, and accessibility processes are prioritized in the mapping, measuring, and managing of AI risks throughout the lifecycle.	GV-3.1	Decision-making related to mapping, measuring, and managing AI risks throughout the lifecycle is informed by a diverse team.		B.4.6; B.5.4	A.4; A.5	Not covered by NIST AI 600-1 §3	2 - Not relevant at this scale	Presupposes a team to diversify. Not present in this operating model; revisit if the operating model gains staff or partners.	-	N/A
	GOVERN 3		GV-3.2	Policies and procedures are in place to define and differentiate roles and responsibilities for human-AI configurations and oversight of AI systems.	7.2	B.3.2; B.4.6; B.5.3; B.6.1.3; B.9.3	A.3; A.4; A.5; A.6; A.9	CBRN; HarmBias; Human-AI; InfoSec	3 - Relevant and inherently met	Human-AI configuration roles are explicit and enforced: HITL gates the harness cannot pass without human action; agents propose, human decides, deterministic core computes.	Spec s12.4, s17.4, s1.2 (boundary), s10.2.	N/A
	GOVERN 4	Organizational teams are committed to a culture that considers and communicates AI risk.	GV-4.1	Organizational policies and practices are in place to foster a critical thinking and safety-first mindset in the design, development, deployment, and uses of AI systems to minimize potential negative impacts.		B.5.2; B.5.4; B.6.1.2; B.6.1.3; B.9.2; B.9.3; B.10.3	A.5; A.6; A.9; A.10	CBRN; Confab; ValueChain	3 - Relevant and inherently met	Safety-first structure, not culture-by-policy: separate-vendor QA, halt/flag, layered self-check, and an explicit push-back norm in the working agreement.	Spec s11, s12; working agreement (push back when warranted).	N/A
GOVERN	GOVERN 4		GV-4.2	Organizational teams document the risks and potential impacts of the AI technology they design, develop, deploy, evaluate, and use, and they communicate about the impacts more broadly.	6.1.4; 7.4	B.5.4; B.5.5; B.8.5	A.5; A.8	Dangerous; Human-AI; IP; Obscene; ValueChain	3 - Relevant and inherently met	Risk/impact documentation and communication are designed surfaces: this self-assessment (a transparency document), and on every assessment the briefing note, public value account, inputs register, CBA spreadsheet and QA report, presented as one set to be read together (each downloaded separately); the evidence lineage is on the QA tab.	This workbook; spec s13, s18.1; the results page's document set.	N/A
GOVERN	GOVERN 4		GV-4.3	Organizational practices are in place to enable AI testing, identification of incidents, and information sharing.		B.6.1.2; B.6.1.3; B.6.2.4; B.6.2.6; B.6.2.7; B.8.2; B.8.3; B.8.4; B.8.5	A.6; A.8	DataPriv; InfoInteg; InfoSec	3 - Relevant and inherently met	Testing/evals infrastructure, a structured observation log of every model and tool action in a run, error surfacing (never silent) - testing and incident identification are built in.	Spec s19, s20, s16, s12.3, s12.6.	N/A
GOVERN	GOVERN 5	Processes are in place for robust engagement with relevant AI actors.	GV-5.1	Organizational policies and practices are in place to collect, consider, prioritize, and integrate feedback from those external to the team that developed or deployed the AI system regarding the potential individual and societal impacts related to AI risks.		B.5.3; B.5.4; B.8.3; B.10.4	A.5; A.8; A.10	Confab; HarmBias; Human-AI	2 - Not relevant at this scale	Policy apparatus for external-feedback collection presupposes organisational scale. The operational mechanism (a user feedback channel) is tracked as the category-4 item at MS-3.3.	See MS-3.3.	N/A
GOVERN	GOVERN 5		GV-5.2	Mechanisms are established to enable the team that developed or deployed AI systems to regularly incorporate adjudicated feedback from relevant AI actors into system design and implementation.		B.5.4; B.5.5; B.6.1.3; B.6.2.6; B.8.3; B.10.4	A.5; A.6; A.8; A.10	Not covered by NIST AI 600-1 §3	2 - Not relevant at this scale	As GV-5.1: regular adjudicated-feedback incorporation presupposes a feedback volume and org structure not present. Builder reviews any feedback directly at this scale.	See MS-3.3.	N/A
GOVERN	GOVERN 6	Policies and procedures are in place to address AI risks and benefits arising from third-party software and data and other supply chain issues.	GV-6.1	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.		B.10.2; B.10.3	A.10	DataPriv; HarmBias; Human-AI; InfoInteg; InfoSec; IP; ValueChain	3 - Relevant and inherently met	Third-party surface is small and known: Anthropic and Google (AI models), Railway (hosting, including the volume that holds cases and runs), open-source libraries and the local embedding model (all-MiniLM-L6-v2, baked into the image), public data sources in the curated source registry, and web sources drafted when a case is created (admitted only at the user's framing sign-off). Every cited source carries a licence/usage note, checked on every run by the QA layer and printed in the inputs register; APIs under standard commercial terms.	Spec s5.2; the as-built dependency inventory; spec s2 (the planned stack, June 2026); the Client Data Handling and Privacy Policy v1.0 s3 (Railway volume); QA tier-1 check (licence note per cited source); the assessment's inputs register (publisher and licence per source).	N/A
GOVERN	GOVERN 6		GV-6.2	Contingency processes are in place to handle failures or incidents in third-party data or AI systems deemed to be high-risk.		B.10.2; B.10.3	A.10	DataPriv; HarmBias; Human-AI; InfoInteg; InfoSec; IP; ValueChain	3 - Relevant and inherently met	Contingency specced: graceful degradation (QA unavailable blocks hand-off - never silently skipped), bounded retries, budget caps, funding-cap hard backstop.	Spec s12.6, s21 (decision record 16).	N/A
MAP	MAP 1	Context is established and understood.	MP-1.1	Intended purposes, potentially beneficial uses, context-specific laws, norms and expectations, and prospective settings in which the AI system will be deployed are understood and documented.	6.1.4	B.5.2; B.5.3; B.5.4; B.5.5	A.5	CBRN; Dangerous; DataPriv; HarmBias; Human-AI; IP; Obscene	3 - Relevant and inherently met	Intended purpose, audience (public bodies, agencies, development banks, third-sector organisations and the consultancies serving them), deployment setting (a hosted system reached through private links, with a recorded public demonstration), and out-of-scope uses documented; the tool's boundary is stated to the user at intake before a case is created.	scoping document s3.1.1 (audience); spec Purpose statement and s1 (governing principle); the AI policy and legal note v1.1 s3 (nature of outputs), s5 (two operating settings), s6 (deployment); the Client Data Handling and Privacy Policy v1.0 s1; intake capability statement (shown at intake). Spec s10.1a, s12.2(a) and specification appendix A (s3.1.1, s3.1.3) record the original two-case scope, superseded by the open user-defined case (build record, module 18).	N/A
MAP	MAP 1		MP-1.2	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.	7.2	B.4.6	A.4	HarmBias; Human-AI	2 - Not relevant at this scale	Demographically diverse interdisciplinary team presupposes scale. Domain (economics) + technical competence present in the builder; revisit if the operating model gains staff or partners.	-	N/A
MAP	MAP 1		MP-1.3	The organization's mission and relevant goals for AI technology are understood and documented.	4.1; 6.2; 7.3; 7.4; 7.5.3			Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Mission and AI goals documented: the build programme purpose (welfare/business/financial value translation) and this build's role in it.	scoping document s1-2; spec purpose statement.	N/A
MAP	MAP 1		MP-1.4	The business value or context of business use has been clearly defined or – in the case of assessing existing AI systems – re-evaluated.	4.1; 5.1	B.2.2; B.5.2; B.6.2.2; B.9.4	A.2; A.5; A.6; A.9	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Business value defined: standardised, auditable evidence of the public value a resilience action creates, delivered faster than a bespoke study; general tool first, cases as data; built to hand off to business- and finance-value analysis.	Spec s1 (governing principle), s18 header; scoping document.	N/A

NIST AI RMF Self-Assessment: cascadiaen												
Function	Category ID	Category Description	Sub ID	Subcategory Text (NIST AI 100-1)	ISO 42001 Clause Refs	ISO 42001 Annex A Controls (B.x.x)	ISO 42001 Annex A Category	GAI Risk Tags (NIST AI 600-1 §3) See Summary for ref.	Triage category	Notes	Evidence reference	Addressed
MAP	MAP 1		MP-1.5	Organizational risk tolerances are determined and documented.	6.1.1			Not covered by NIST AI 600-1 §3	4 - Relevant and required	Organisational risk tolerance is operationalised (halt/flag thresholds) but never stated as a documented tolerance. One short risk-tolerance statement; pairs with GV-1.3. Addressed 2026-07-02: risk tolerances determined and documented per dimension (note s4), user-set framing: advisory tool not critically embedded; imprecision accepted when labelled; invisibility not tolerated; transparency and record-keeping as the primary risk tools. Reviewed 2026-09-16 (v1.1): tolerances unchanged; s4 restated for client engagements.	Partial: spec s11.4. Gap: documented tolerance statement. ADDRESSED 2026-07-02: the AI policy and legal note s4 (risk tolerance). REVIEWED 2026-09-16: the AI policy and legal note v1.1 s4.	Yes
MAP	MAP 1		MP-1.6	System requirements (e.g., 'the system shall respect the privacy of its users') are elicited from and understood by relevant AI actors. Design decisions take socio-technical implications into account to address AI risks.		B.5.4; B.5.5; B.6.2.2	A.5; A.6	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	The technical spec is the elicited-requirements artefact; socio-technical implications addressed as binding constraints (real/dummy separation, HITL, transparency).	Spec throughout; s5.6, s18.1 (binding UI constraints), s9A.	N/A
MAP	MAP 2	Categorization of the AI system is performed.	MP-2.1	The specific tasks and methods used to implement the tasks that the AI system will support are defined (e.g., classifiers, generative models, recommenders).		B.4.2; B.4.3; B.4.4; B.4.5; B.4.6; B.6.2.3	A.4; A.6	DataPriv; InfoInteg; IP	3 - Relevant and inherently met	Tasks and methods precisely defined: deterministic CBA core + five agent roles (case authoring, orchestration, research, scenario, write-up) + separate-vendor QA, with the deterministic/probabilistic boundary as the load-bearing line.	Spec s1.1-1.2, s7-s10. Case-authoring agent (added after the spec's four roles).	N/A
MAP	MAP 2		MP-2.2	Information about the AI system's knowledge limits and how system output may be utilized and overseen by humans is documented.		B.6.2.7; B.8.2; B.9.3	A.6; A.8; A.9	InfoInteg; ValueChain	3 - Relevant and inherently met	Knowledge limits are first-class: confidence never defaulted, Indicative flags, qualitative components never numeric, research bounded per run; human oversight of output at HITL gates + QA.	Spec s8.3 (guards), s12.1 (per-run budget), s11, s11.3 (limitations caveat check).	N/A
MAP	MAP 2		MP-2.3	Scientific integrity and 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.		B.6.1.3; B.6.2.4; B.6.2.7; B.7.2; B.7.3; B.7.4; B.7.5; B.7.6	A.6; A.7	HarmBias; InfoInteg; InfoSec; IP	3 - Relevant and inherently met	Scientific integrity/TEVV designed in: four-dimension benefit-transfer match drives confidence, provenance mandatory on every parameter, golden/failure eval datasets.	Spec s4.7, s5, s19.2; acceptance-criteria doc.	N/A
MAP	MAP 3	AI capabilities, targeted usage, goals, and expected benefits and costs compared with appropriate benchmarks are understood.	MP-3.1	Potential benefits of intended AI system functionality and performance are examined and documented.		B.5.2; B.5.3; B.5.4; B.5.5	A.5	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Potential benefits examined and documented in the scoping/spec (welfare-value visibility for resilience decisions; general tool demonstrated on cases).	scoping document s1-2; spec s1, specification appendix A.	N/A
MAP	MAP 3		MP-3.2	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.	8.2; 8.3; 8.4	B.5.2; B.5.3; B.5.4; B.5.5	A.5	Not covered by NIST AI 600-1 §3	4 - Relevant and required	Mitigations for AI errors are extensive, but the harm analysis itself (what a wrong/misleading valuation could cost, over-trust in outputs) is not documented. One impact & residual-risk note covers MP-3.2 + MP-5.1 + MG-1.4. Addressed 2026-07-02: AI-error classes and their costs documented (note s5.2) - chiefly non-monetary/reputational at MVP scale, escalating at commercial scale; tied to the risk-tolerance statement. Reviewed 2026-09-16 (v1.1): s5.2 restated for client engagements - error classes translate to client mis-decisions and credibility costs; reputational only on the public demonstration.	Mitigations: spec s11, s12. Gap: documented error cost/impact analysis. ADDRESSED 2026-07-02: the AI policy and legal note s5 (impact & residual risk). REVIEWED 2026-09-16: the AI policy and legal note v1.1 s5.2.	Yes
MAP	MAP 3		MP-3.3	Targeted application scope is specified and documented based on the system's capability, established context, and AI system categorization.	4.3	B.5.2; B.5.3; B.5.4; B.5.5	A.5	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Application scope tightly specified: valuation components and methods are those on the TEV reference table; components the table does not quantify are acknowledged in words, never monetised; no primary research (e.g. new stated-preference studies). The boundary is stated to the user at intake, and a case outside it is refused with the reason.	TEV reference table; intake capability statement (shown at intake). Spec s10.1a, s12.2(a), s6.4 and scoping s3.1.3 record the original two-case boundary, superseded by the open user-defined case (build record, module 18).	N/A
MAP	MAP 3		MP-3.4	Processes for operator and practitioner proficiency with AI system performance and trustworthiness – and relevant technical standards and certifications – are defined, assessed, and documented.	7.2	B.4.6	A.4	Dangerous; HarmBias; Human-AI; InfoInteg	2 - Not relevant at this scale	Operator/practitioner proficiency processes presuppose a workforce of operators. Assessments are run by the builder (domain expert) or by named holders of private access links; the workflow is guided (wayfinding, rules shown before a figure is entered, refusals explained with a recovery action) and every gate decision is recorded against the run. Revisit if clients operate the tool in volume.	Spec s18.1 (wayfinding).	N/A
MAP	MAP 3		MP-3.5	Processes for human oversight are defined, assessed, and documented in accordance with organizational policies from the GOVERN function.		B.6.1.3; B.6.2.7; B.8.2	A.6; A.8	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Human oversight defined, documented, and enforced outside the model: eight-step workflow with explicit HITL gates the runner cannot pass.	Spec s12.2 (the four gates (a)-(d) and QA-halt decisions, gate lettering corrected 2026-07-21), s17.2, s12.4.	N/A
MAP	MAP 4	Risks and benefits are mapped for all components of the AI system including third-party software and data.	MP-4.1	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.	4.1	B.2.2; B.9.2; B.9.4	A.2; A.9	CBRN; Dangerous; DataPriv; HarmBias; InfoSec; IP; Obscene; ValueChain	3 - Relevant and inherently met	Component and legal risk mapping in place: full third-party inventory, a licence/usage note per cited source (populated, and printed in the inputs register), citation-linked extracts; standard API terms. Client figures are used on the client's confirmation that they are entitled to share them (companion policy s2).	The as-built dependency inventory; spec s5.2, s14.1 (attribution). The assessment's inputs register. The Client Data Handling and Privacy Policy v1.0 s2.	N/A
MAP	MAP 4		MP-4.2	Internal risk controls for components of the AI system, including third-party AI technologies, are identified and documented.		B.6.2.7; B.8.2; B.10.3	A.6; A.8; A.10	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Internal risk controls per component identified and documented: every agent tool call passes the harness pre-flight (schema/provenance/permissions/budget/guards); model calls are budget-checked; server-side web searches during case authoring are logged and capped per request, not pre-flighted; QA re-verifies post-hoc.	Spec s12.1, s12.5, s11.	N/A

NIST AI RMF Self-Assessment: cascadiaen												
Function	Category ID	Category Description	Sub ID	Subcategory Text (NIST AI 100-1)	ISO 42001 Clause Refs	ISO 42001 Annex A Controls (B.x.x)	ISO 42001 Annex A Category	GAI Risk Tags (NIST AI 600-1 §3) See Summary for ref.	Triage category	Notes	Evidence reference	Addressed
MAP	MAP 5	Impacts to individuals, groups, communities, organizations, and society are characterized.	MP-5.1	Likelihood and magnitude of each identified impact (both potentially beneficial and harmful) based on expected use, past uses of AI systems in similar contexts, public incident reports, feedback from those external to the team that developed or deployed the AI system, or other data are identified and documented.	6.1.2	B.5.2	A.5	CBRN; Dangerous; HarmBias; Human-AI; InfoInteg; InfoSec; Obscene	4 - Relevant and required	Impact likelihood/magnitude (beneficial and harmful) not yet characterised - who could rely on outputs, what over-trust or misquoted figures could cause. Part of the single impact & residual-risk note (with MP-3.2, MG-1.4). Addressed 2026-07-02: beneficial and harmful impacts identified with qualitative likelihood/magnitude and mitigations (note s5.1); MVP-vs-commercial scale asymmetry stated up front. Reviewed 2026-09-16 (v1.1): s5.1 re-rated for client engagements (reliance on outputs is the purpose of an engagement) and exposure of client material added, mitigated under the companion policy.	Gap: impact characterisation. Mitigations exist: spec s11, s13. ADDRESSED 2026-07-02: the AI policy and legal note s5 (impact & residual risk). REVIEWED 2026-09-16: the AI policy and legal note v1.1 s5.1; the Client Data Handling and Privacy Policy v1.0 (the engagement folder: system revision 16, item 28).	Yes
	MAP 5		MP-5.2	Practices and personnel for supporting regular engagement with relevant AI actors and integrating feedback about positive, negative, and unanticipated impacts are in place and documented.		B.5.3; B.6.1.3; B.6.2.6; B.8.3	A.5; A.6; A.8	Human-AI; ValueChain	2 - Not relevant at this scale	Regular structured engagement with external AI actors presupposes scale. Feedback channel (MS-3.3) is the scale-appropriate mechanism.	See MS-3.3.	N/A
MEASURE	MEASURE 1	Appropriate methods and metrics are identified and applied.	MS-1.1	Approaches and metrics for measurement of AI risks enumerated during the MAP function are selected for implementation starting with the most significant AI risks. The risks or trustworthiness characteristics that will not – or cannot – be measured are properly documented.	6.1.1; 6.1.2			CBRN; HarmBias; Human-AI; InfoInteg; Obscene	3 - Relevant and inherently met	Measurement is risk-prioritised by design: the core friction (methodology fidelity) carries the deepest eval treatment; harder-to-measure agent quality addressed via failure-mode set; fuzziness explicitly acknowledged.	Spec s19.1-19.2; acceptance-criteria doc; s9A.	N/A
MEASURE	MEASURE 1		MS-1.2	Appropriateness of AI metrics and effectiveness of existing controls are regularly assessed and updated, including reports of errors and potential impacts on affected communities.		B.5.2; B.5.4; B.5.5; B.6.2.4	A.5; A.6	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Metrics and controls are iterated, not static: evals loop, acceptance criteria revisited at testing, prompt changes eval-gated, human reviewer maintains golden set.	Spec s19.1, s10.4, s11.2 (human review tier).	N/A
MEASURE	MEASURE 1		MS-1.3	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.	6.1.2; 9.2.2	B.5.2; B.5.4; B.5.5	A.5	CBRN; DataPriv; HarmBias; Human-AI	2 - Not relevant at this scale	Independent human assessors presuppose org scale. The different-vendor QA agent (Gemini vs Claude - not checking its own homework) is the scale-appropriate structural analogue, noted here without claiming equivalence to human independent review.	Spec s11.1 (decision record 1).	N/A
MEASURE	MEASURE 2	AI systems are evaluated for trustworthy characteristics.	MS-2.1	Test sets, metrics, and details about the tools used during TEVV are documented.		B.4.2; B.6.2.4; B.6.2.7; B.8.4	A.4; A.6; A.8	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	TEVV assets documented: golden, failure-mode, and retrieval datasets; suite table with method and cost; Inspect AI tooling; pass*k defined.	Spec s19.2-19.4, s20.	N/A
MEASURE	MEASURE 2		MS-2.2	Evaluations involving human subjects meet applicable requirements (including human subject protection) and are representative of the relevant population.		B.6.2.4	A.6	Dangerous; DataPriv; HarmBias; Human-AI; InfoInteg; InfoSec	1 - Not relevant for this build type	No human-subjects evaluation in this build type (economic valuation tool; no human-subject testing or population-representative studies).	-	N/A
MEASURE	MEASURE 2		MS-2.3	AI system performance or assurance criteria are measured qualitatively or quantitatively and demonstrated for conditions similar to deployment setting(s). Measures are documented.		B.6.2.6; B.7.4	A.6; A.7	CBRN; Confab; Dangerous; DataPriv; HarmBias; InfoInteg; InfoSec	3 - Relevant and inherently met	Performance measured for deployment-like conditions: end-to-end integration runs, whole-run QA on every run, record-and-replay agent evals over real traces.	Spec s19.3, s20, s11.2.	N/A
MEASURE	MEASURE 2		MS-2.4	The functionality and behavior of the AI system and its components – as identified in the MAP function – are monitored when in production.	9.1	B.6.2.6; B.6.2.8	A.6	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Production monitoring built in: every live run writes a structured observation log of each model and tool action and is QA'd in full (no sampling at this scale); errors logged, never swallowed; an operator run log lists every run. OpenTelemetry instrumentation is in place, but its trace export is not enabled on the hosted deployment.	Spec s16, s11.2, s12.3, s12.6; operator console run log; improvement backlog item FI-013 (trace export).	N/A
MEASURE	MEASURE 2		MS-2.5	The AI system to be deployed is demonstrated to be valid and reliable. Limitations of the generalizability beyond the conditions under which the technology was developed are documented.		B.6.2.4; B.6.2.5; B.6.2.7; B.8.2	A.6; A.8	Confab; Dangerous; Human-AI; InfoInteg; InfoSec	3 - Relevant and inherently met	Validity/reliability demonstrated via the free deterministic golden tier + pass*k on critical cases; generalizability limits explicit (stated tool boundary, confidence framework, worked-examples-non-binding).	Spec s19.4, s9A, intake capability statement (shown at intake); spec s10.1a records the original two-case limit, superseded by the open user-defined case (build record, module 18).	N/A
MEASURE	MEASURE 2		MS-2.6	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.	8.2	B.6.2.4; B.6.2.6; B.6.2.8	A.6	CBRN; Confab; Dangerous; DataPriv; HarmBias; Human-AI; InfoInteg; InfoSec; IP; Obscene; ValueChain	3 - Relevant and inherently met	Fail-safe by construction: blocking harness guards, HALT criteria stop the workflow for the user, budget/iteration caps, QA-unavailable blocks hand-off, funding-cap backstop.	Spec s11.4, s12.1, s12.6, s21.	N/A
MEASURE	MEASURE 2		MS-2.7	AI system security and resilience – as identified in the MAP function – are evaluated and documented.		B.2.3; B.3.2; B.5.2; B.6.1.2; B.6.2.3; B.7.2; B.9.3	A.2; A.3; A.5; A.6; A.7; A.9	Dangerous; DataPriv; HarmBias; Human-AI; InfoInteg; InfoSec; ValueChain	4 - Relevant and required	Resilience is specced (s12.6) and access/cost controls exist (decision record 16), but no documented security review of the threat surface: public light gate, free-text user input into agents (prompt injection), MCP exposure, key handling. Requires a short security note with mitigations. Addressed 2026-07-02: security threat-surface review written into spec s21 - five surfaces (public access, prompt injection, MCP incl. access-key gating + tool tiering with HITL preserved, key handling, data/supply chain), each with mitigation and residual; commercial-shapes forward look. Updated 2026-09 (v1.1): live access is through individually killable, expirable, usage-capped private links issued from an operator console; the public face serves recorded assessments only, with no model spend; the access-code gate is off. Retrieved web content remains within the prompt-injection surface (spec s21: injection can corrupt proposals, never outcomes).	Partial: spec s21 (decision record 16), s12.1 (permissions), s12.6. Gap: documented threat-surface review. ADDRESSED 2026-07-02: spec s21 security threat-surface review (5 surfaces incl. substantive MCP treatment; commercial-shapes forward look). UPDATED 2026-09-16: the AI policy and legal note v1.1 s6 (security of the deployed system: private links, operator console, client material, prompt injection, machine interface (MCP), keys and cost, resilience), which supersedes the spec s21 review of the earlier shared-code design (it carries forward that review's access, prompt-injection, MCP and key-handling surfaces as deployed).	Yes

NIST AI RMF Self-Assessment: cascadiaen												
Function	Category ID	Category Description	Sub ID	Subcategory Text (NIST AI 100-1)	ISO 42001 Clause Refs	ISO 42001 Annex A Controls (B.x.x)	ISO 42001 Annex A Category	GAI Risk Tags (NIST AI 600-1 §3) See Summary for ref.	Triage category	Notes	Evidence reference	Addressed
MEASURE	MEASURE 2		MS-2.8	Risks associated with transparency and accountability – as identified in the MAP function – are examined and documented.	6.1.2	B.5.4; B.5.5; B.6.1.2; B.7.2; B.9.3	A.5; A.6; A.7; A.9	HarmBias; Human-AI; InfoInteg; IP	3 - Relevant and inherently met	Transparency/accountability is the design centre: methodology trace as architectural primitive, per-parameter provenance, QA always visible, overrides logged. On every assessment the inputs register gives each value in force with its source, and the QA report lists every finding with any override as issue, signatory's decision and justification verbatim, flagged for manual review.	Spec s13, s5-5.6, s18.1, s11.4; the assessment's inputs register and QA report.	N/A
MEASURE	MEASURE 2		MS-2.9	The AI model is explained, validated, and documented, and AI system output is interpreted within its context – as identified in the MAP function – to inform responsible use and governance.		B.6.2.5; B.6.2.7; B.7.5; B.8.2	A.6; A.7; A.8	Confab; HarmBias; InfoInteg	3 - Relevant and inherently met	Explainability designed in: every value carries its source and the rationale for where it was set (evidence cards, inputs register), the account shows where the value comes from by component, and the briefing note states the method and the confidence of each value in plain language; progressive disclosure on the results page.	Spec s13, s8.1; the assessment's inputs register, results page and briefing note.	N/A
MEASURE	MEASURE 2		MS-2.10	Privacy risk of the AI system – as identified in the MAP function – is examined and documented.		B.2.3; B.5.2; B.7.2; B.7.3	A.2; A.5; A.7	HarmBias; Human-AI; InfoInteg; IP	4 - Relevant and required	No personal data is sought: the system does not ask for it, clients are asked not to include it, and any included is deleted on request (system revision 16, item 28); no user accounts; access is by private link. Client material is held in one folder per engagement, kept 6 months from delivery of the package (unless otherwise agreed) and then deleted whole, leaving one content-free log line (system revision 16, item 28). AI vendors are used on paid tiers whose terms exclude training use; their retention is stated in the policy. Addressed 2026-09-16: Client Data Handling and Privacy Policy authored; the engagement folder, hold-period dates, deletion (at the end of the hold period and on request) and the content-free log line are built by system revision 16, item 28.	The Client Data Handling and Privacy Policy v1.0 (s2-s8, s11); spec s17.3; paid billing on both AI accounts (the AI policy and legal note v1.1 s3, Third-party AI services); implemented in system revision 16, item 28.	Yes
MEASURE	MEASURE 2		MS-2.11	Fairness and bias – as identified in the MAP function – are evaluated and results are documented.		B.5.4; B.5.5	A.5	Dangerous; Environ; HarmBias	3 - Relevant and inherently met	No individual-level decisions; bias risk here = representativeness of transferred valuation evidence, handled directly by the four-dimension benefit-transfer match with confidence never defaulted, and contested evidence capped Indicative.	Spec s8.3, s4.2 (ConfidenceLevel), s6.2 (contested-evidence handling).	N/A
MEASURE	MEASURE 2		MS-2.12	Environmental impact and sustainability of AI model training and management activities – as identified in the MAP function – are assessed and documented.		B.4.5; B.5.5	A.4; A.5	Dangerous; Environ	3 - Relevant and inherently met	No model training; inference only, bounded by harness-enforced spend ceilings per run and per case authoring, with spend recorded per run; scenario recomputation uses no model calls. Footprint de minimis at this scale.	Spec s10.1 (cost), s12.1 (budget), s16; per-run spend in the run record's meter; the operator console run log shows paid-call counts per run; its authoring log shows dollar spend per case authoring.	N/A
MEASURE	MEASURE 2		MS-2.13	Effectiveness of the employed TEVV metrics and processes in the MEASURE function are evaluated and documented.		B.6.2.4; B.6.2.6	A.6	Confab; HarmBias; InfoInteg	3 - Relevant and inherently met	TEVV effectiveness is itself evaluated: QA meta-eval (does QA catch planted failure traces), pass*k reliability on critical cases, retrieval-quality eval rather than assumed retrieval.	Spec s19.3 (QA meta-eval), s19.4, s14.1.	N/A
MEASURE	MEASURE 3	Mechanisms for tracking identified AI risks over time are in place.	MS-3.1	Approaches, personnel, and documentation are in place to regularly identify and track existing, unanticipated, and emergent AI risks based on factors such as intended and actual performance in deployed contexts.	4.4; 8.2; 8.4			Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Risk tracking scale-appropriate: structured observation log per run, build record as the running record, periodic review cadence (at each material change to the system or its operating model (for example a new capability, a change of hosting or AI provider, or the first client engagement) and annually).	Spec s12.3, s16; build record; Lifecycle Map stage 5-6.	N/A
MEASURE	MEASURE 3		MS-3.2	Risk tracking approaches are considered for settings where AI risks are difficult to assess using currently available measurement techniques or where metrics are not yet available.	10.1	B.6.2.6; B.6.2.8	A.6	Confab; Human-AI	3 - Relevant and inherently met	Hard-to-measure risks explicitly handled: agent-layer 'correct is fuzzier' acknowledged, failure-mode dataset + LLM-judge + flags where exact metrics don't exist.	Spec s19.2, s11.4 (FLAG set).	N/A
MEASURE	MEASURE 3		MS-3.3	Feedback processes for end users and impacted communities to report problems and appeal system outcomes are established and integrated into AI system evaluation metrics.		B.8.2; B.8.3; B.8.4	A.8	HarmBias; Human-AI; InfoInteg	4 - Relevant and required	No end-user feedback/problem-reporting channel in the spec. Public deployment needs a light contact/feedback route (and note on how reports are handled). Cheap, signals well. Addressed 2026-07-02: feedback route specced as a binding frontend requirement (spec s18.1) pointing to builds@cascadiaen.com with handling stated; reported failure modes feed the failure-mode test dataset (spec s19.2). Channel is built with the frontend. Updated 2026-09-16 (v1.1): the channel is email - engagements@cascadiaen.com (published on the website for engagements) and builds@cascadiaen.com (named in the tool's own messages), both to the builder. Handling: every report is answered; a report that reveals a defect is fixed and added to the failure-mode test dataset. The in-page feedback control of the first interface was not carried into the current pages; the server route remains unused.	Gap: feedback channel on deployed surface. Covers GV-5.x operational intent. ADDRESSED 2026-07-02: specced as binding requirement, spec s18.1 (feedback route -> builds@cascadiaen.com; handling stated) + s19.2 (reported failure modes feed the failure-mode test dataset). Channel exists at build. UPDATED 2026-09-16: the two published email addresses (website; the tool's messages). Spec s18.1's in-page feedback link and handling statement are superseded by the two published email addresses; handling is as stated in K. Adding a report to the failure-mode test dataset is a manual operator step.	Yes
MEASURE	MEASURE 4	Feedback about efficacy of measurement is gathered and assessed.	MS-4.1	Measurement approaches for identifying AI risks are connected to deployment context(s) and informed through consultation with domain experts and other end users. Approaches are documented.	9.1	B.5.4; B.5.5; B.6.2.4	A.5; A.6	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Measurement approaches are domain-expert-owned; the acceptance-criteria document is user-owned ('writing evals is product management'); human reviewer maintains the golden set.	Acceptance-criteria doc; spec s19.1, s11.2.	N/A
MEASURE	MEASURE 4		MS-4.2	Measurement results regarding AI system trustworthiness in deployment context(s) and across the AI lifecycle are informed by input from domain experts and relevant AI actors to validate whether the system is performing consistently as intended. Results are documented.	9.1; 9.2.1	B.8.2; B.8.3	A.8	Confab; Human-AI; InfoInteg; InfoSec	3 - Relevant and inherently met	Every run is human-verified at the HITL gates; each decision, adjustment and QA override (with the signatory's justification verbatim) is recorded in the run record and the exported QA report; run state + trace persist results.	Spec s17.2-17.3, s11.3, s13. The assessment's QA report.	N/A
MEASURE	MEASURE 4		MS-4.3	Measurable performance improvements or declines based on consultations with relevant AI actors, including affected communities, and field data about context-relevant risks and trustworthiness characteristics are identified and documented.	9.3.1	B.6.2.6; B.6.2.7	A.6	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Performance changes tracked through the eval suites: prompt/agent changes are eval-gated, regressions visible against golden/failure sets; field data = per-run observation logs.	Spec s10.4, s19.3, s16.	N/A

NIST AI RMF Self-Assessment: cascadiaen												
Function	Category ID	Category Description	Sub ID	Subcategory Text (NIST AI 100-1)	ISO 42001 Clause Refs	ISO 42001 Annex A Controls (B.x.x)	ISO 42001 Annex A Category	GAI Risk Tags (NIST AI 600-1 §3) See Summary for ref.	Triage category	Notes	Evidence reference	Addressed
MANAGE	MANAGE 1	AI risks based on assessments and other analytical output from the MAP and MEASURE functions are prioritized, responded to, and managed.	MG-1.1	A determination is made as to whether the AI system achieves its intended purposes and stated objectives and whether its development or deployment should proceed.		B.6.1.3; B.6.2.4; B.9.2; B.9.3; B.9.4	A.6; A.9	Not covered by NIST AI 600-1 §3	4 - Relevant and required	Staged gates exist (scope -> spec -> this assessment -> build plan) but the deployment go/no-go determination is not yet defined as an artefact. Requires a short deployment decision record (go/no-go + residual risks) at deploy time. Addressed 2026-07-02: deployment decision record specced as binding (spec s21) - publish checklist covering intended-purpose judgement, tests/QA, residual risks accepted, transparency surfaces, governance artefacts, cost protections (funding cap + token/budget limits + gate), unpublish ability confirmed, and the signed decision; executed as a dated build record entry at deploy. Publishing without it is out of process. Executed 2026-07-31: deployment decision record written (GO), with accepted residual risks enumerated. Clarified 2026-09-16 (v1.1): the 2026-07-31 GO covered the beta deployment; the system has been tested and iterated since. A go/no-go for client engagements is made and recorded before the first engagement, on the same checklist.	Process: working agreement sequence; specification appendix B. Gap: deployment decision record (Lifecycle Map stage 4), ADDRESSED 2026-07-02: specced as binding requirement, spec s21 deployment decision record (dated build record entry: tests/evaluations passed, QA verified, residual risks accepted, transparency surfaces working, governance artefacts in place, cost protections incl. token/budget limits tested, publish decision). Written at deploy; publishing without it is out of process. EXECUTED 2026-07-31 (build record, deployment decision record).	Yes
MANAGE	MANAGE 1		MG-1.2	Treatment of documented AI risks is prioritized based on impact, likelihood, and available resources or methods.	6.1.2; 6.1.3; 6.1.4; 9.3.3			Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Risk treatment explicitly prioritised: core friction gets deepest treatment; halt (enforcement) vs flag (information) severity split; breadth-vs-depth watch-item named.	Spec s11.4, s9A, specification appendix D (residual risks).	N/A
MANAGE	MANAGE 1		MG-1.3	Responses to the AI risks deemed high priority, as identified by the MAP function, are developed, planned, and documented. Risk response options can include mitigating, transferring, avoiding, or accepting.	6.1.1; 6.1.2; 6.1.3; 6.1.4			Human-AI; InfoSec	3 - Relevant and inherently met	Responses to the priority risk (methodology infidelity) are the architecture itself: deterministic boundary, harness enforcement, separate-vendor QA, evals; documented throughout.	Spec s1.2, s12, s11, s19.	N/A
MANAGE	MANAGE 1		MG-1.4	Negative residual risks (defined as the sum of all unmitigated risks) to both downstream acquirers of AI systems and end users are documented.		B.5.3; B.5.4; B.6.2.7; B.8.2	A.5; A.6; A.8	Not covered by NIST AI 600-1 §3	4 - Relevant and required	Residual risks to users and the downstream build (the downstream build consumes the hand-off package) not yet documented. Part of the single impact & residual-risk note (with MP-3.2, MP-5.1); include what QA/harness do NOT catch. Addressed 2026-07-02: six-item residual-risk register documented (note s5.3) - what the mitigations do not catch, incl. the downstream hand-off to business-value analysis; the register is the source of the outputs' limitations caveat. Reviewed 2026-09-16 (v1.1): a seventh residual added (client material held by the AI providers, outside cascadiaen's deletion); the run's limitations caveat covers this register's categories for the run's own caveats (the register is not an input to the run) and is shown on the briefing note with the standard limitations statement, and in the QA report (system revision 16, item 29).	Gap: residual-risk statement. Hand-off surface: spec s4.5 (hand-off package). ADDRESSED 2026-07-02: the AI policy and legal note s5 (impact & residual risk). REVIEWED 2026-09-16: the AI policy and legal note v1.1 s5.3; the Client Data Handling and Privacy Policy v1.0 s3, s10.	Yes
MANAGE	MANAGE 2	Strategies to maximize AI benefits and minimize negative impacts are planned, prepared, implemented, documented, and informed by input from relevant AI actors.	MG-2.1	Resources required to manage AI risks are taken into account – along with viable non-AI alternative systems, approaches, or methods – to reduce the magnitude or likelihood of potential impacts.	7.1	B.4.2	A.4	Not covered by NIST AI 600-1 §3	3 - Relevant and inherently met	Non-AI alternatives applied wherever viable by principle: all computation deterministic (no LLM), LLMs only where genuine NL reasoning is needed, templated/structured UI elsewhere; resources bounded by budget caps.	Spec s10.1 (cost discipline), s1.2, s12.1.	N/A
MANAGE	MANAGE 2		MG-2.2	Mechanisms are in place and applied to sustain the value of deployed AI systems.	7.1; 10.1	B.3.3; B.6.1.2; B.6.1.3; B.6.2.4; B.6.2.6; B.7.2	A.3; A.6; A.7	CBRN; Confab; Dangerous; DataPriv; HarmBias; Human-AI; InfoInteg; InfoSec; IP; Obscene	3 - Relevant and inherently met	Value sustained scale-appropriately: pinned versions, config-swappable model IDs, durable run state, monitoring; periodic review cadence.	Spec s2, s11.1 (configurable), s17.3, s16.	N/A
MANAGE	MANAGE 2		MG-2.3	Procedures are followed to respond to and recover from a previously unknown risk when it is identified.	6.1.1; 6.1.2; 6.1.3; 10.2			ValueChain	3 - Relevant and inherently met	Unknown-risk response: HALT stops the workflow for user decision; operational recovery specced separately (never mistaken for methodology failure); ability to unpublish/disable.	Spec s11.4, s12.6, s21.	N/A
MANAGE	MANAGE 2		MG-2.4	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.		B.6.1.3; B.6.2.7; B.8.2; B.9.4	A.6; A.8; A.9	Human-AI; InfoSec	3 - Relevant and inherently met	Override/deactivate mechanisms real and assigned: HITL overrides logged and visible, harness halts, private access links individually killable/expirable/usage-capped, service disable = builder action, funding cap = hard stop.	Spec s11.4 (overrides logged), the AI policy and legal note v1.1 s6 (deployed access controls); Lifecycle Map stage 4.	N/A
MANAGE	MANAGE 3	AI risks and benefits from third-party entities are managed.	MG-3.1	AI risks and benefits from third-party resources are regularly monitored, and risk controls are applied and documented.		B.10.2; B.10.3	A.10	CBRN; DataPriv; HarmBias; InfoInteg; InfoSec; IP; ValueChain	3 - Relevant and inherently met	Third-party risks monitored at review cadence: model updates, dataset vintages (source registry carries vintage/retrieval date), API terms; model IDs swappable via config; a vendor change is a code change (a new transport).	Spec s5.2, s2; Lifecycle Map stage 6.	N/A
MANAGE	MANAGE 3		MG-3.2	Pre-trained models which are used for development are monitored as part of AI system regular monitoring and maintenance.		B.4.4; B.6.2.6	A.4; A.6	CBRN; Confab; Dangerous; DataPriv; HarmBias; Human-AI; InfoInteg; IP; Obscene; ValueChain	3 - Relevant and inherently met	Pre-trained models = Claude/Gemini APIs: model IDs pinned in config, changes deliberate, behaviour re-checked against eval suites when models change. The QA judge currently runs on a preview model endpoint, to be moved to a generally available one when it exists.	Spec s2, s10.1, s19.3.	N/A
MANAGE	MANAGE 4	Risk treatments, including response and recovery, and communication plans for the identified and measured AI risks are documented and monitored regularly.	MG-4.1	Post-deployment AI system monitoring plans are implemented, including mechanisms for capturing and evaluating input from users and other relevant AI actors, appeal and override, decommissioning, incident response, recovery, and change management.	9.2.1	B.6.2.6; B.8.3; B.10.4	A.6; A.8; A.10	CBRN; Confab; HarmBias; Human-AI; InfoInteg; InfoSec	3 - Relevant and inherently met	Post-deployment monitoring plan exists at scale: per-run observation log + whole-run QA + operator run log + periodic output review + link kill/expiry or service disable; change management via versioned artefacts. The user-feedback/appeal element is the MS-3.3 category-4 item.	Spec s16, s11.2, s1.5 (versioned artefacts); operator console; MS-3.3.	N/A
MANAGE	MANAGE 4		MG-4.2	Measurable activities for continual improvements are integrated into AI system updates and include regular engagement with interested parties, including relevant AI actors.	9.3.3	B.6.2.4; B.6.2.6	A.6	Dangerous; HarmBias; Human-AI	3 - Relevant and inherently met	Continual improvement integrated: evals loop iterates agent + eval together, acceptance criteria revisited at testing, self-assessment versioned (v1.1, reviewed September 2026) with review cadence.	Spec s19.1, specification appendix B; this workbook version line.	N/A

NIST AI RMF Self-Assessment: cascadiaen												
Function	Category ID	Category Description	Sub ID	Subcategory Text (NIST AI 100-1)	ISO 42001 Clause Refs	ISO 42001 Annex A Controls (B.x.x)	ISO 42001 Annex A Category	GAI Risk Tags (NIST AI 600-1 §3) See Summary for ref.	Triage category	Notes	Evidence reference	Addressed
MANAGE	MANAGE 4		MG-4.3	Incidents and errors are communicated to relevant AI actors, including affected communities. Processes for tracking, responding to, and recovering from incidents and errors are followed and documented.	9.3.2	B.6.2.6; B.8.5	A.6; A.8	Confab; DataPriv; InfoInteg; InfoSec	3 - Relevant and inherently met	Incidents tracked and responded to via observation log + error surfacing (never silent) + halt screens in-run; link kill (one or all) or service disable available; a site notice needs a page edit and redeploy. No affected-community apparatus needed at this scale.	Spec s12.3, s12.6, s18.1 (halt screens).	N/A

NIST AI RMF: Build Lifecycle Mapping (cascadiaen Interpretation)			
CAVEAT: NIST AI RMF is deliberately lifecycle-agnostic. This mapping is a cascadiaen interpretation for sequencing build activities, not an official NIST artefact. Present it as such.			
Lifecycle Stage	NIST Categories	Key Activities / Questions	Notes for cascadiaen Builds
1: Plan & Design	GOVERN 1–2 (setup) MAP 1 (context) MAP 2 (categorisation) MAP 3 (benefits/costs) MAP 4.1 (component risk inventory)	What is this build for and who uses it? What are the legal/regulatory constraints? What is the risk tolerance? What are the intended and non-intended uses? What components (third-party models, data, APIs) are in scope? Document: purpose, scope, risk tolerance, component inventory.	For each build: complete MAP 1.1, 1.3–1.5 before architecture decisions are finalised. GOVERN setup applies across all builds: set once, reference per build.
2: Collect & Process Data	MAP 4 (third-party data/software) MEASURE 2.10 (privacy) MEASURE 2.11 (bias in data) MEASURE 2.12 (environmental) MAP 2.3 (TEVV / data quality)	What data sources are used and under what terms? Are there privacy risks in the data pipeline? Are there representativeness or bias risks in reference data? What is the environmental cost of model calls? Document: data provenance, licence terms, known limitations.	Primarily relevant for models trained on proprietary data. For builds using third-party LLMs (Claude and Gemini), focus on input/output data handling. Still relevant for the curated source registry, web sources drafted when a case is created, and client-supplied figures.
3: Build & Validate	MEASURE 1 (metrics) MEASURE 2.1–2.5 (TEVV) MEASURE 2.6 (safety) MEASURE 2.7 (security/resilience) MEASURE 2.8 (transparency/accountability) MEASURE 2.9 (explainability) MEASURE 2.13 (TEVV efficacy)	What does 'working correctly' look like for this build? How is output validated against known benchmarks? Can the system fail gracefully and within known limits? Is the reasoning chain traceable and explainable? Are security boundaries clearly defined? Document: test criteria, QA agent check logic, known limitations.	The QA layer (separate LLM, evidence lineage, halt/flag logic) directly addresses MEASURE 2.1–2.9 and 2.13. Document the mapping explicitly in the assessment.
4: Deploy	MAP 5 (impact characterisation) MANAGE 1 (go/no-go) MANAGE 2.1 (resources) MANAGE 2.4 (override/deactivate) MEASURE 2.5 (validity at deployment)	Has a go/no-go determination been made? Are impacts to users and third parties characterised? Are mechanisms in place to override or deactivate if needed? Are resource constraints documented? Document: deployment decision record, known residual risks.	For cascadiaen systems: 'deploy' = releasing the hosted assessment system (a public demonstration face and private access links). Go/no-go = builder determination, recorded in a dated deployment decision record. Override/deactivate = kill, expire or cap an access link, or disable the service. Residual risk documentation is part of the transparency document.
5: Operate & Monitor	MEASURE 2.4 (production monitoring) MEASURE 3 (risk tracking) MANAGE 2.2 (sustain value) MANAGE 2.3 (unknown risk response) MANAGE 4 (post-deployment) MEASURE 3.3 (feedback loops)	How is system behaviour monitored post-deployment? How are unanticipated risks identified and escalated? Is there a feedback mechanism for users? Is there an incident response process? Document: monitoring approach, feedback channels, incident log.	Lightest function at this operating scale. Monitoring = periodic review of build outputs. Incident response = ability to take build offline. Some subcategories here will fall into category 2 (not relevant at this scale).
6: Continuous / Cross-cutting	ALL of GOVERN MEASURE 4 (measurement feedback) MANAGE 3 (third-party) MANAGE 2.3 (unknown risk) Improvement loops	Is the governance structure holding up as builds evolve? Are third-party components (Claude API, MCP, data sources) still fit for purpose? Are measurement approaches being refined? Is the self-assessment itself being updated? Document: version log, review cadence, third-party change log.	GOVERN is cross-cutting by design. For cascadiaen: periodic review at each material change to the system or its operating model (for example a new capability, a change of hosting or AI provider, or the first client engagement) and annually. Third-party monitoring includes Claude and Gemini model and endpoint changes (the QA judge runs on a preview endpoint), the embedding model, hosting changes, MCP SDK changes, source registry and dataset vintage updates.

ISO 42001 Annex A: 9-Category Aggregation (Reference Guide)

Source: ISO/IEC 42001 Annex A (B.x.x controls), aggregated into 9 categories. Cross-reference sourced from NIST's published crosswalk. cascadiaen is not claiming ISO 42001 alignment or certification.

Annex A Category	Controls Included	Relevance to cascadiaen Builds
A.2: Policies	B.2.2 AI policy; B.2.3 Alignment with other org policies; B.2.4 Review of AI policy	Core governance artefact for cascadiaen. The explicit AI policy is the AI policy & legal note (s2), recorded at GV-1.2; client data handling is set in the Client Data Handling and Privacy Policy.
A.3: Internal Organisation	B.3.2 AI roles and responsibilities; B.3.3 Reporting of concerns	At this operating scale, roles collapse to the builder. Roles and decision rights are documented in the working agreement's decision protocol and the AI policy & legal note s2 (Governance and accountability).
A.4: Resources	B.4.2 Resource documentation; B.4.3 Data resources; B.4.4 Tooling; B.4.5 System/computing; B.4.6 Human resources	Relevant: data sources (curated source registry; web sources drafted per case), tooling (Claude and Gemini APIs, MCP), computing (AI vendor infrastructure and the application host). Component inventory (NIST MAP 4) covers this.
A.5: Impact Assessment	B.5.2 AI system impact assessment process; B.5.3 Documentation; B.5.4 Assessing impacts on individuals/groups; B.5.5 Assessing societal impacts	High relevance. The build programme's TEV methodology IS an impact assessment framework. The self-assessment should surface this mapping explicitly.
A.6: Lifecycle	B.6.1.x Responsible development objectives/processes; B.6.2.x Operation, monitoring, V&V, documentation, deployment, event logs	Most operationally relevant category. QA agent, evidence lineage, halt/flag logic, and technical documentation all map here.
A.7: Data	B.7.2 Data for development/enhancement; B.7.3 Acquisition; B.7.4 Quality; B.7.5 Provenance; B.7.6 Preparation	Relevant for benefit-transfer and exposure evidence in the source registry, web sources drafted per case, client-supplied figures, and data passed to AI agents. Provenance is mandatory per parameter and printed in the inputs register; client material is handled under the Client Data Handling and Privacy Policy (MS-2.10).
A.8: Information for Interested Parties	B.8.2 System documentation/info for users; B.8.3 External reporting; B.8.4 Communication of incidents; B.8.5 Information for interested parties	The transparency document (this self-assessment) and the documents every assessment produces (briefing note, public value account, inputs register, CBA spreadsheet, QA report) and the Client Data Handling and Privacy Policy (what a client is told about their material) directly respond to this category. High relevance.
A.9: Use	B.9.2 Processes for responsible use; B.9.3 Objectives for responsible use; B.9.4 Intended use of AI system	Intended use is documented in build programme scoping (s3.1.1), as extended by the open user-defined case (build record, module 18) and the AI policy & legal note v1.1 s3, satisfying B.9.4. Responsible use objectives map to the build programme's methodology visibility principle.
A.10: Suppliers/Customers	B.10.2 Allocating responsibilities; B.10.3 Suppliers; B.10.4 Customers	Suppliers = Anthropic and Google (AI models), Railway (hosting), open-source libraries including the MCP SDK and the local embedding model, data sources. Customers = organisations commissioning assessments and holders of private access links; visitors to the public demonstration. Client material is handled under the Client Data Handling and Privacy Policy. Third-party risk (NIST MAP 4 / MANAGE 3) covers suppliers.