AI Governance Policy — Layer 1 (Foundation)

GENERATED from 10-rules-en — do not edit by hand

Company-wide policy — applies to every employee, every department, at all times. Principle: the level of control rises with the level of risk. Below are the minimum rules (red lines) for the Foundation tier (L1).

Identity & approved tools

AUTH-001 — Use only company-approved AI tools

tiers: L1, L2, L3, L4

Only AI tools **under a company contract whose input is excluded from training** may be used for company work. **Personal AI accounts are strictly forbidden** for any company data ("Shadow AI", L1 §8).

Current list (`30-registry/ai-tools.yml`):

- Google Workspace Gemini (company)
- Claude (Team / Enterprise)

AUTH-002 — Every tool must sign in with an @vitalify.asia account

tiers: L1, L2, L3, L4

Every tool used with AI — AI tools (`AUTH-001`) and work tools / connectors — must authenticate with an `@vitalify.asia` account; personal accounts are forbidden — IT cannot revoke access when an employee leaves, data ownership cannot be enforced, and there is no audit log for incident investigation.

Current list (`30-registry/connectors-mcp.yml`):

- Figma
- Google Docs/Sheets/Slides/Drive
- Draw.io

AUTH-003 — Confirm the tool is company-managed and the data is allowed before input

tiers: L1, L2, L3, L4

Before entering any **non-public** information, confirm **both**: the tool is company-managed (contracted, input excluded from training) **and** the data is permitted under policy. If either is missing → stop and ask PMO. *Paid does not mean safe.*

AUTH-004 — Adding a new AI tool must go through the approval flow

tiers: L1, L2, L3, L4

Do not adopt a new AI tool on your own. Approval flow: staff proposes → direct manager → CoE consolidates → **Dev Director** decides → CoE communicates back. Until approved, treat the tool as **not permitted**.

Data allowed / not allowed into AI

DATA-001 — Never put client data / production secrets into AI

tiers: L1, L2, L3, L4

Never enter: production/test credentials (IDs, passwords, API keys), end-user personal data / clients' raw data (test logs, transaction histories), and top-secret internal information (undisclosed financials, M&A). This is the company's **red line #1**. At L2 it is partly enforced through the `deny` list in settings.json (Read/Write of secret-file patterns).

DATA-002 — Three-level data classification

tiers: L1, L2, L3, L4

Level	May I enter it?
Top secret (credentials, end-user PII, M&A)	❌ No — even on company tools (see `DATA-001`)
Internal business data (minutes, proposals, specs)	⚠️ Yes — **company tools only**. Proper nouns such as client names may be entered if normally shared internally; for project-specific exceptions, ask your manager
General information (public, lookups, proofreading of non-confidential text)	✅ Yes — no restriction

DATA-003 — Anonymize by sensitivity level (not everything)

tiers: L1, L2, L3, L4

COE decision (#2): follow L1, relaxing L2's anonymization rule. Not every client name must be anonymized. Specifically:

- Client names / proper nouns **may be entered as-is** if normally shared internally (e.g. over email/chat).
- RFPs / specs / proposals **may be uploaded** after removing secrets, PII and internal IP (see `DATA-004`) — no need to replace with placeholders.
- Only **sensitive commercial figures** (price, rate, budget, margin) and top-secret data under `DATA-001`/`DATA-002` **must be anonymized**.

DATA-004 — Document upload procedure (RFP / spec / design)

tiers: L1, L2, L3, L4

Before uploading a document into AI: **(1)** confirm the contract allows AI use for it — if unsure, ask the client or don't use it; **(2)** remove passwords, API keys, internal IP addresses, personal data; **(3)**

use an approved tool only. Review sensitive cases with PMO.

DATA-007 — AI-assisted code must be IP-clean

tiers: L1, L2, L3, L4

AI-generated code that enters a deliverable must be **reviewed and scanned for license/originality** (to avoid reproducing GPL/AGPL code that would "contaminate" the IP) and must meet the same quality & security bar as human-written code. The client's IP-ownership terms apply to AI code exactly as to all other code.

DATA-008 — No generative AI for commercial brand assets

tiers: L1, L2, L3, L4

Do not use generative AI for final brand designs delivered to clients or used as our own identity (logo, mascot, trademark) — AI may reproduce protected work undetectably and cannot be verified "clean". AI may be used for inspiration/moodboards/drafts; the final asset is made by a human designer. If a client requests it, escalate to PMO.

DATA-011 — Send work content only to the approved AI tool

tiers: L1, L2, L3, L4

Do not send source code, documents, data or logs to any external service other than the approved AI tool you are working in — no pastebins, online formatters or converters, other AI/LLM APIs, webhook testers or public file-sharing links. When several AI tools work on the same content, each of them must be approved (`AUTH-001`). Agents fetch only from approved domains (`RULES-006`).

Safe AI tool configuration

RULES-007 — AI must not act on production, company or client systems

tiers: L1, L2, L3, L4

AI must not take autonomous action on company systems, client systems or the Production environment: no sending emails, calling APIs or operating cloud systems without human approval. AI must never directly access, operate, modify, deploy to, or execute commands against Production — including deployment scripts and CI/CD pipelines, configuration files and environment variables, credentials/secrets/tokens/API keys, infrastructure, databases, cloud resources and operational systems. Even when Production scripts, credentials or configurations exist in the source code or repository, AI may only **review them and suggest changes**; it must not execute, apply, test or use them against Production. In particular, AI must not: run Production deployment or migration scripts; execute commands using Production credentials; modify Production infrastructure, databases, configurations or secrets; disable or bypass Production security controls; trigger CI/CD jobs targeting Production. All Production actions must be reviewed, approved and **manually performed** by an authorized human.

Incident response

RES-001 — Report a security incident immediately

tiers: L1, L2, L3, L4

On discovering an incident (data leak, secret/credential pushed to GitHub, client data exposed, suspicious access…): **report immediately** to your direct manager **and** CoE at the same time, in the `[SECURITY] Incident Report` channel, as soon as you notice — don't wait to confirm details, don't fix it yourself first. Give a short description: what, where (repo/tool/system), when you noticed.

RES-002 — On an incident: stop the task, revoke access, trace the impact

tiers: L1, L2, L3, L4

Beyond reporting, the response must include: **stop** the running task/agent, **revoke** the access/credentials of the agent/person involved, and **trace** the impact scope back to the person/command that triggered it. L1 §11 currently has only "report to manager + CoE" — it **lacks** the stop/revoke/trace steps Kato asked to add.