Data Meaning

Control Tower for Claude: security brief

For your CISO, IT security and privacy teams · Version 1.0.4

SECURITY BRIEF

Control Tower for Claude

Sections 1 to 2

1. What it is

Control Tower for Claude shows who uses Claude, for what work, at what cost, and whether its settings match your policy. It runs in your Azure subscription and changes nothing in Claude unless a person confirms it.

2. Where it runs

It installs from its Microsoft Marketplace listing, which passed Microsoft's certification review, into a resource group you choose. The template gives Data Meaning no access to your subscription. If you want us to run the install, your IT team gives a Data Meaning consultant access to that resource group for the install and removes it afterward. It creates:

- one App Service web app serving the pages and, at `/mcp`, the MCP server (HTTPS only, TLS 1.2);
- a function app for the daily run at 14:00 UTC and a 30-minute check for requests for more usage;
- PostgreSQL with public access disabled, reached through a private endpoint; the apps' connection strings require TLS;
- a storage account with a private HR-file container;
- Key Vault with role-based access, soft delete and purge protection;
- Log Analytics and Application Insights.

The web app is public, behind company sign-in. Key Vault accepts only Entra ID sign-in with an Azure role. The storage account also accepts its account keys, so anyone allowed to list them, such as a Contributor on the resource group, can read the HR file. Only the database is private.

3. What it reads

From Anthropic. The reading code sends only GET requests.

- **Enterprise key (required)**: `read:analytics`, `read:compliance_activities`, `read:compliance_user_data`, `read:compliance_org_data`, `read:members`, `read:spend_limits`. `read:compliance_user_data` can read every chat, file and project; of that content, the code reads only chat and session text, to label it. Only the function app, which runs the daily run and the 30-minute check, holds this key. Create it without `delete:compliance_user_data`; the checks flag any key in the organization with a delete scope.
- **Admin API key (optional)**: spend on apps built on the Claude API, and key expiry. The code only reads with this key.
- **Labeling key (optional)**, from a Claude Console organization, separate from your Enterprise organization and possibly under a different Anthropic agreement. Labeling runs under its terms and is billed to it.

From you:

- **The HR file**, a CSV in the storage container: email, department and manager email; optionally sub-department, location, role, hire date and cost center.
- **A Microsoft 365 shared mailbox**, off until configured. It sends through Microsoft Graph with `Mail.Send` and `MailTips.ReadBasic.All`, limited to that mailbox by Exchange RBAC for Applications, and checks recipients' automatic-reply status; it reads no mailbox contents.
- **Azure Reader role**, off until configured. With it, the daily run reads who changed its resource group and who holds roles there.

SECURITY BRIEF

Control Tower for Claude

Sections 4 to 5

4. What it keeps, and what it never keeps

Kept: usage and cost per person per day; members (email, name, Claude role, activity dates); each HR file; each chat's or session's date, person, product, model, use case, confidence and screened description; a 16-character hash of each session's opening, to spot automated runs; settings, roles and keys; project names, owners, and chat and file counts; Anthropic's activity-feed events; the access log and activity list; the text of Claude chat questions it could not answer; and records of features you turn on: requests for more usage and spend limits, license and access-review decisions, emails sent, policy acknowledgments, evidence packs and record samples, and, with Reader, who holds roles on the resource group, and who changed its apps or role assignments.

Never kept: chat or session text, file contents, IP addresses, device and browser details, request bodies, and the chat and file titles and file names in activity events (project names are kept, from the project list). No table has a column for chat text, and tests check that labeling writes none.

5. How chats are labeled

Each new chat, and each Claude Code, Cowork and Office add-in session, is:

1. read through the Compliance API: by default at most 12,000 characters; for a session, its first page of turns, with tool inputs and outputs cut to one byte;
2. sent to Claude with your use-case list, using the labeling key; the answer can only be a listed use case, or "other";
3. discarded. The text is held only in memory.

A second call, with the same text, writes one sentence describing the kind of task. A pattern check rejects any digit and the characters @ $ € £ %. A further Claude call reads only the sentence and rejects anything that could identify a person, company, client, product, project or place, any number or date, personal, employment or sensitive business matters, and instructions. HR and personal conversations get no sentence. Each sentence is stored with its chat's record but shown only by use case, never beside a person.

Control Tower for Claude

Sections 5 to 7

To check accuracy, a stronger model re-reads a random sample; only the sample size and the share judged right are stored. Your Legal team should confirm that the Console organization's agreement permits sending chat text for labeling.

6. Who sees what

- **Admins:** every page, including named lists (unused licenses, broadest users, top spenders), and the admin console.
- **Leaders:** department and company totals and every department's bill. Their one named list is the access log on the Governance page.
- **Managers**, from the HR file: their own reporting line as one group, without names or Governance.
- **Members**, everyone else with a license: their own My Claude pages.
- **Stewards**, chosen by an admin: the Actions page for their assigned departments and kinds of action, where they can confirm those changes. Stewards exist only once the change key is set up.
- **An outside reviewer you invite**, as a guest in your Entra tenant: read-only access, until a date you set at most 180 days ahead, to approved evidence packs and record lists for one period, including per-person records a sample needs. Never the dashboards, My Claude or Claude chat.

Only admins, leaders and managers can ask from Claude chat, with the same limits. A Data Meaning consultant can sign in only with an account in your Entra tenant, seeing what its role allows.

Small-group rule. No group under 5 people is shown on its own, anywhere. You can raise the minimum, never lower it. Exceptions: admins' named lists, the access log (admins and leaders), each person's own pages, a department's own bill (seen only by admins, leaders and that department's manager), and per-person records an outside reviewer sees in a sample.

7. What it can change in Claude

These features are off at install. They switch on only when your Primary Owner creates a separate change key and your IT team stores it in Key Vault, readable by the web app alone. Without it, the app has no way to make changes in Claude. The key carries `read:members`, `write:members`, `read:spend_limits` and `write:spend_limits`.

With it, an admin, or a steward for their departments, can confirm removing people who left or have not used Claude in 60 days, withdrawing invitations pending 14 days or more, answering requests for more usage (a denial needs a reason), and setting monthly spend limits. Only admins invite people. A person can give back their own license from their own page; it is removed at once, without an admin confirming.

The server checks permission and a session-bound token on every change. For each person it re-reads Anthropic, acts, reads back and records; a batch is at most 25 people and stops at the first error from Anthropic. Email links only open the app, after sign-in. People with an owner or admin role in Claude cannot be removed from the app. Where your identity provider manages membership, the app shows its steps instead.

It can never delete chats, files or projects, change Claude's settings, create keys, or act without a person.

8. Sign-in, secrets and logging

Sign-in is OpenID Connect against your Entra ID tenant, with PKCE and checks of each token's signature, issuer, audience and expiry. People are matched to roles by the email address in their sign-in token. Sessions are signed, Secure, HttpOnly cookies lasting 12 hours. On App Service, where the Marketplace template runs it, the web app will not start without sign-in configured.

Secrets go from the install screens into Key Vault. Each app's identity gets one role per secret it needs: the web app no Anthropic read or labeling key, the daily run no sign-in secret. The apps use a database login that cannot change the schema; the web app also holds the database administrator login, which it uses to update the schema each time it starts. The daily run can read and write the storage account.

Control Tower for Claude

Sections 8 to 10

Logging. The access log records every page view, every answer given in Claude chat, and the text of any question it could not answer. The activity list records every change and who made it. The app's own login cannot change or delete these logs; the database administrator login can. Two days after each day ends, that day's rows in the access log, the activity list and the activity feed are fingerprinted with SHA-256 and chained to the day before, so a later change to a fingerprinted day shows whenever the chain is rechecked, which happens each time an evidence pack is built and on demand. Someone with that login could rebuild the whole chain, so each fingerprint can also go to a locked storage container you create, off until configured. When a labeling run fails, its record keeps only the error type or Anthropic's error message, never chat text.

9. Data leaving your subscription

1. Reads, and changes a person confirms, to Anthropic's APIs.
2. Labeling, description, screening and accuracy calls to Claude under the Console organization; all but screening carry chat or session text.
3. Answers to Claude chat questions, into that person's conversation, and so into their Claude chat history.
4. Sign-in with your identity provider, and emails, if configured, through Microsoft Graph.
5. Monthly search terms for kinds of work (never people or chat text) to skills.sh and GitHub, and public skill files to Claude, for My Claude tips.
6. Vanta or Drata, only if credentials are added and a person presses Send: an approved evidence pack. An admin can also turn on a daily feed of license removals for people who left, identified by a hash of their email; it also needs those credentials.
7. Within Azure, with Reader: reads of the resource group's Activity Log and role assignments.
8. Typefaces from Google Fonts, fetched by viewers' browsers.

10. Retention, and when you stop

Data is kept while the service runs. A chat deleted before labeling is never read; a label and its description stay when the chat is later deleted. Database backups are kept 7 days and Azure's Log Analytics logs 30. The access log and activity list are kept for the life of the service.

SECURITY BRIEF

Control Tower for Claude

Sections 10 to 12

When the contract ends, we export the full access log for you. Revoke the Enterprise key and any change key in Claude's organization settings, and the Admin and labeling keys in the Claude Console. If you have given us access, we delete the resource group within 30 days and confirm in writing; otherwise we give your team the steps. A deleted Key Vault keeps its secrets recoverable for 30 days and cannot be purged sooner, so revoke the keys first. A locked fingerprint container holds only dates, row counts and fingerprints, and its storage account and resource group cannot be deleted until its policy ends, so the 30-day deletion does not cover them. Text sent for labeling is held by Anthropic under the Console organization's terms; emails stay in mailboxes, and anything sent to Vanta or Drata stays there.

11. What we ask of you before go-live

1. **Legal approval** of this brief, of chat labeling, and of the Console organization's terms.
2. **An employee notice**, from our template, sent before labeling starts.
3. **The readiness checklist:** the install (your IT team, or our consultant with temporary resource-group access), the keys, an Entra app registration, the HR file, and your admins and leaders.

12. Questions to ask us

Ask us to walk through the template, each key, one labeled chat's stored record, or deletion: info@datameaning.com