Izzet.Labs

FREE ENTERPRISE GUIDE · IZZET LABS

# Claude in the Enterprise

## A Security-First Implementation Guide

**Surfaces | Governance | Evaluation | Rollout**

For IT leaders, security teams & decision-makers bringing Claude into their organization · izzetlabs.com

| Document Type | Audience | Last Updated |
|---|---|---|
| Implementation Guide | IT Leaders \| Security \| Decision-Makers | June 2026 |

Most organizations already have Claude inside them — often before IT or security has signed off. This is the operator's playbook for getting ahead of it: the right surface for each job, data and access under control, and a safe path to roll it out across the company.

# Table of Contents

INTRODUCTION

# Why an Operator Wrote This

A field guide from the security and engineering side of the table

This guide was written by a practitioner who deploys and governs technology inside enterprise environments for a living — not by a vendor selling you the product. That perspective matters, because most AI adoption advice is written to make adoption frictionless. The security reality is messier.

Claude is one of the most capable AI platforms available, and it is already inside most organizations — pasted into a browser tab, running in a developer's terminal, or quietly summarizing documents on someone's desktop. The question for IT and security teams is no longer **whether** Claude is in use. It is whether it is being used on the right surface, on the right plan, with the right controls. Get those three things right and Claude becomes a governed, auditable, high-leverage tool. Get them wrong and you have unmanaged data exposure with a friendly interface.

## Who this guide is for

- **IT and security leaders** deciding how — and whether — to bring Claude into the organization.
- **Decision-makers** who need a clear, vendor-neutral picture of the data, identity, and compliance implications.
- **Practitioners** tasked with writing the policy, running the pilot, and standing up the controls.

## How to use it

Part 1 maps the five surfaces Claude shows up on and helps you match each to the right job. Part 2 is the security and governance foundation — data handling, identity, logging, and compliance. Part 3 gives you a repeatable way to evaluate the platform before you commit. Part 4 is a phased rollout playbook. Part 5 is the quick-reference set you can hand to a team: cheat sheet, checklist, and a vendor question list.

| | |
|---|---|
| NOTE | **Independent and vendor-neutral.** Izzet Labs is an independent consultancy and is **not affiliated with, endorsed by, or sponsored by Anthropic**. Every product fact in this guide is drawn from Anthropic's public documentation as of June 2026 and is cited at the end. Products change quickly — verify the current details against the linked sources before you make decisions. |
| WATCH OUT | This guide is general information for planning purposes, not legal, compliance, or security advice for your specific situation. Regulatory obligations vary by industry and jurisdiction. Confirm your requirements with qualified counsel and your own security team before deploying any AI tool on regulated or sensitive data. |

PART 1

# The Claude Surface Map

Five ways Claude shows up at work — and why the choice is a security decision

## 1.1 Why “Which Surface” Is a Security Question

“Claude” is not one product. It is a family of surfaces that share a model but differ enormously in what they can touch. A chat tab can only see what a person pastes into it. An agentic desktop session can read a folder of files, open applications, and act in a browser. A coding agent can run shell commands against a repository and CI secrets. The same underlying intelligence carries a very different blast radius depending on where it runs.

That is why surface selection is the first security decision, not an afterthought. Before you write a single policy line, you need a shared map of the surfaces, what each one can access, and which jobs belong on which surface.

## 1.2 The Five Surfaces

| Surface | What it is |
|---|---|
| **Claude apps (Chat)** | The web, desktop (macOS/Windows), and mobile chat experience. Conversational; sees only what the user provides. The default entry point for most employees. |
| **Claude Cowork** | An agentic workspace inside the desktop app for non-developers. Given access to a folder, it reads, drafts, and organizes files, and can use connectors or control the screen/browser to finish a task. Available on paid plans. |
| **Claude Code** | A command-line and IDE coding agent for developers. Combines repository files, shell commands, tool calls, and CI context in a single session. Also runs on the web in isolated VMs and headless in CI/CD. |
| **Claude API / Platform** | Programmatic access for building Claude into your own products and internal tools. Available directly from Anthropic and via Amazon Bedrock, Google Vertex AI, and Microsoft Foundry. |
| **Agentic surfaces** | Browser and computer-use capabilities (e.g., Claude in Chrome, desktop control) that let Claude act in live applications. The most powerful — and the largest attack surface. |

**NOTE** Claude apps, Cowork, and Claude Code now live side by side in the same desktop application. Convenient for users — but it means a single device can reach three very different capability levels. Govern at the surface and plan level, not just "Claude: yes/no."

## 1.3 What Data Each Surface Touches

The attack surface is defined by reach. Map each surface to what it can see and do, and your data-exposure picture becomes concrete:

| Surface | Data it can reach | Exposure |
|---|---|---|
| **Claude apps (Chat)** | Pasted text, uploaded files, and connected data sources the user enables. Exposure = whatever a person chooses to put in. | Low–Medium |
| **Claude Cowork** | A selected local folder, connected apps (e.g., Slack, calendar), and — when used — on-screen content and the browser. Local sessions stay on-device. | Medium–High |
| **Claude Code** | Repository source, local files, shell command output, environment variables, and CI secrets in non-interactive runs. | High |
| **API / Platform** | Exactly what your application sends — which can include customer data at scale. Governed by your code and key configuration. | Depends on design |
| **Agentic surfaces** | Live web pages and application state, including anything visible while acting. Vulnerable to instructions hidden in the content it reads. | High |

**WATCH OUT** Agentic surfaces introduce a risk text chat does not: **prompt injection**. A malicious instruction hidden in a web page, document, or email can try to redirect Claude into actions the user never intended. Anthropic runs automated detection for this, but it is an evolving threat — treat any surface that acts on untrusted content as higher-risk, and keep a human in the loop for consequential actions.

## 1.4 Which Surface for Which Job

| The job | Use this surface | Why |
|---|---|---|
| Quick questions, drafting, brainstorming | Claude apps (Chat) | Lowest reach for the most common need. |

| | | |
|---|---|---|
| File-heavy knowledge work for non-developers | Claude Cowork | Scoped to a folder; finishes whole tasks, not single prompts. |
| Working in a codebase | Claude Code | Built for repositories, with permission controls for shell actions. |
| Embedding Claude in a product or internal tool | API / Platform | You control data flow, retention, and deployment region. |
| Acting across live web apps | Agentic surfaces | Powerful; reserve for vetted workflows with oversight. |

**BEST PRACTICE** Default employees to the lowest-reach surface that does the job. Most needs are met by Chat on a commercial plan. Escalate to Cowork, Code, or agentic surfaces deliberately, per use case — not by default.

## 1.5 Plans: The Line That Actually Matters

From a security standpoint, the most important distinction is not Free versus paid — it is **consumer terms versus commercial terms**. They carry different default data-handling commitments, covered in detail in Part 2.

| Plan | Terms | What it means for you |
|---|---|---|
| Free / Pro / Max | Consumer | Individual plans. Governed by Consumer Terms. Data-training behavior depends on a user setting (see 2.2). Not for regulated data. |
| Team / Enterprise (“Claude for Work”) | Commercial | Organization plans. Governed by Commercial Terms. Inputs/outputs are not used for training by default; add SSO, SCIM, audit logging, and ZDR at the Enterprise level. |
| Claude for Education | Commercial | Tailored for institutions; commercial data-handling terms apply. |
| Claude for Government | Commercial | Government-tailored offering with its own terms and posture. |

**KEY IDEA** If an employee is using a personal **Free, Pro, or Max** account for company work, that activity sits under **consumer terms** and outside your organization's controls — no SSO, no central audit trail, and data-training may be switched on. Moving company use onto a **commercial** plan is the single highest-leverage governance step you can take.

PART 2

# Security & Governance Foundation

Data, identity, logging, and your real compliance posture

## 2.1 The Shared-Responsibility Model

Cloud security taught everyone a useful mental model, and it applies cleanly to AI platforms. The provider secures the platform; you govern how your organization uses it. Anthropic is responsible for the security of the model and infrastructure, its certifications, and the data-handling commitments in your contract. **You** are responsible for who has access, what data your people put in, which surfaces are allowed, and how usage is logged and reviewed.

| Responsibility | What it covers |
|---|---|
| Anthropic secures | Model and infrastructure security; encryption in transit and at rest; certifications and audits; honoring your contractual data-handling terms; isolation of cloud code sessions. |
| You govern | Identity and access; acceptable-use policy; which surfaces/plans are permitted; what data classes are allowed; monitoring and audit-log review; vendor due diligence; incident response. |

**KEY IDEA** No certification, setting, or contract removes your half of the work. The platform can be perfectly secure and you can still leak data through an ungoverned personal account. Plan for both halves.

## 2.2 Data Handling: Training & Retention

This is the question security teams ask first: is our data used to train the model, and how long is it kept? The answer depends entirely on whether you are on commercial or consumer terms.

| Account type | Training & retention (default behavior) |
|---|---|
| Commercial (Team, Enterprise, API, Bedrock/Vertex, Government) | **Not used to train** generative models by default. Training only happens if an admin explicitly opts in (e.g., a development partner program on the first-party API). Standard retention is limited (around 30 days for Team/Enterprise and Claude Code), and Zero Data Retention is available for qualifying customers. |

| | |
|---|---|
| Consumer (Free, Pro, Max) | Training depends on a **user setting**. If a user allows their data to be used to improve Claude, retention extends to **five years**. If they decline, retention is roughly **30 days**. The choice lives in the user's privacy settings and can be changed at any time. |

**WATCH OUT** **The shadow-AI trap.** An employee on a personal Pro or Max account may have model-training switched on — meaning company information they paste could be retained for years and used to improve future models. This is not a flaw in the product; it is the default consumer behavior. The fix is organizational: move company work onto a commercial plan, where training is off by default, and tell employees plainly not to use personal accounts for work.

**NOTE** Settings and exact retention windows change over time and differ by surface and contract. Treat the figures above as the shape of the policy, and confirm the current specifics in the Anthropic Privacy Center and your own agreement before relying on them.

## 2.3 Zero Data Retention (ZDR)

Zero Data Retention is the control most regulated teams care about. Under ZDR, your inputs and outputs are not stored at rest after the response is returned — abuse and safety checks still run, but they happen in-pipeline and nothing persists afterward, except where retention is legally required.

| Aspect | Detail |
|---|---|
| What it does | Prevents conversation content from being written to disk after processing. |
| Who it's for | Teams handling protected health information, financial data, or other regulated categories. |
| How to get it | Available to qualifying Enterprise customers via addendum, and for the API with appropriately configured keys. Coordinate with Anthropic and your legal team. |
| Trade-offs | Some convenience features that rely on stored history (such as session resumption) may be limited. Confirm impacts for your workflows. |

**BEST PRACTICE** ZDR is necessary but not sufficient for regulated data. Pair it with access controls, data-loss prevention, and mandatory human review of any output that touches sensitive records.

## 2.4 Identity & Access Management

Enterprise plans put Claude behind your existing identity stack. These are the controls that turn “everyone has an account” into governed access:

| Control | Guidance |
|---|---|
| Single sign-on (SSO) | Authenticate through your identity provider using SAML 2.0 or OIDC. Centralizes access and enforces your existing login policies. |
| SCIM provisioning | Automate account creation and removal. When someone leaves your IdP, their Claude access is revoked automatically — minutes, not days. |
| Role-based access | Assign administrator and standard roles. Admins control organization settings; everyone else gets usage permissions only. |
| Managed devices | Combine with MDM policy to restrict desktop and CLI use to enrolled, managed devices for higher assurance. |

**BEST PRACTICE** Deprovisioning is where most organizations quietly fail. SCIM closes the gap: tie Claude to your IdP so a departing employee loses access the moment they are offboarded — including any locally cached sessions policy can reach.

## 2.5 Audit Logging & Monitoring

You cannot govern what you cannot see. Enterprise tooling exposes usage and agent activity so it can flow into your existing monitoring stack.

| Capability | Guidance |
|---|---|
| Compliance / audit APIs | Programmatic access to usage and audit data for export into your systems of record. |
| Agent activity streaming | Cowork can stream tool calls, file access, and approval states to your SIEM via OpenTelemetry, so agentic actions are observable, not opaque. |
| SIEM export & retention | Route logs to your SIEM on a schedule. Retain for at least 90 days to support SOC 2, and longer for regulated industries. |
| What to capture | For chat: session events, file uploads, conversation metadata. For agentic/desktop: connector connections, tool invocations, and external service calls. |

**NOTE** Logging matters most for the agentic surfaces, where Claude takes actions rather than just producing text. If an agent can act on your behalf, its actions belong in the same audit trail you hold every other privileged process to.

## 2.6 Your Real Compliance Posture

Anthropic maintains a substantial set of certifications and compliance offerings. Knowing what each one does — and does not — cover keeps your risk assessment honest.

| Framework | What it means |
|---|---|
| SOC 2 Type II | Independent audit of security controls over time. Detailed report available under NDA via the Trust Center; a SOC 3 summary is publicly available. |
| ISO 27001:2022 | Information security management system certification. |
| ISO/IEC 42001:2023 | Management system standard specific to artificial intelligence — relatively rare, and relevant for AI governance. |
| CSA STAR / NIST 800-171 | Additional cloud-security and controlled-information assurance frameworks. |
| HIPAA | Business Associate Agreements offered to qualifying customers; requires ZDR, and some features may be restricted under a BAA. Consumer plans are not HIPAA-eligible. |
| GDPR / data residency | Data-residency options via AWS EU regions and Google Vertex; documentation supporting EU obligations is available in the Trust Center. |

**KEY IDEA** Certifications are a floor, not a ceiling. SOC 2 attests to Anthropic's controls — it does not make **your** deployment compliant. Asking “is Claude SOC 2 compliant?” is like asking whether a document written in a word processor is compliant: the answer depends on how **you** handle the data around it. Layer your own access controls, logging, and review on top.

**NOTE** Anthropic's Trust Center (trust.anthropic.com) is the authoritative source for current certifications, subprocessor lists, and compliance artifacts, and supports document requests under NDA. Start there for any formal vendor review.

## 2.7 Deployment Topology

Where Claude runs shapes where your data travels — a central question for regulated and data-residency-sensitive workloads.

| Path | Summary |
| --- | --- |
| Direct (Anthropic) | Claude apps and the first-party API. Simplest path; data handled under your Anthropic agreement. |
| Amazon Bedrock | Access Claude within your AWS account and region, under your cloud provider's controls and agreements. |
| Google Vertex AI | Access Claude within Google Cloud, supporting regional and private-networking requirements. |
| Microsoft Foundry | Access Claude within the Microsoft cloud ecosystem. |

**BEST PRACTICE** If data residency or VPC isolation is a hard requirement, a cloud-hosted path (Bedrock or Vertex) often fits existing controls better than direct access. Map this to your data-classification policy before choosing.

PART 3

# Evaluate Before You Deploy

A vendor-grade method for trusting any AI tool

## 3.1 Why Structured Evaluation Beats Vibes

AI tools enter organizations bottom-up: an employee finds one useful and starts using it, often on a personal account, long before procurement or security is involved. By the time leadership notices, there are a dozen tools in play and no consistent basis for trusting any of them. A repeatable evaluation method is the antidote — it lets you say yes quickly to the safe choices and no clearly to the rest, using the same criteria every time.

**KEY IDEA** Treat AI tools exactly like any other vendor handling your data. The novelty of the technology is not a reason to skip the due diligence you would apply to a SaaS provider — it is a reason to apply it more carefully.

## 3.2 Questions to Ask Any AI Vendor

Before you trust a tool with company data, get clear answers — from documentation, not sales claims — to each of these:

| Area | Ask |
|---|---|
| Data & training | Is our data used to train models? Is that the default, or opt-in? Does it differ by plan? |
| Retention | How long is data kept? Is Zero Data Retention available? What is the deletion process? |
| Compliance | Which certifications are held (SOC 2, ISO 27001, ISO 42001)? Is a HIPAA BAA available if needed? Are reports available under NDA? |
| Identity | Does it support SSO (SAML/OIDC) and SCIM? What access roles exist? |
| Subprocessors & residency | Who are the subprocessors? Can data be kept in a specific region? Are VPC/private-network options available? |

| | |
|---|---|
| Logging | What audit logging is exposed? Can it stream to our SIEM? Are agent actions captured? |
| Operational | How are incidents disclosed? How are models and versions controlled? What does the contract say about liability and indemnification? |

## 3.3 A Lightweight Scoring Approach

Turn those questions into a score so comparisons are objective and repeatable. Rate each tool across a consistent set of categories — for example data handling, security controls, compliance, vendor trust, operational fit, and cost — then weight the categories by what matters most for your risk profile. A simple, documented rubric beats a gut call you cannot defend in an audit, and it scales: the same scorecard works for the next ten tools your teams want to adopt.



## 3.4 Common Evaluation Mistakes

| Mistake | Why it bites |
|---|---|
| “It's SOC 2, so we're compliant” | A vendor's certification covers the vendor's controls, not your deployment. You still own access, data handling, and review. |
| Ignoring shadow AI | If you only evaluate the tool you're formally adopting, you miss the personal accounts already in use. Sweep for them. |
| Conflating consumer and commercial terms | The same brand can carry very different data commitments by plan. Evaluate the plan you'll actually use. |
| Trusting marketing over docs | Capabilities and policies live in documentation and contracts. Verify there, not in a slide deck. |

| | |
|---|---|
| Treating it as one-and-done | Models, features, and terms change. Re-assess on a schedule and after major provider announcements. |

PART 4

# The Rollout Playbook

Pilot, expand, govern, optimize — without losing control

## 4.1 Phased Rollout Overview

Resist the urge to flip Claude on for everyone at once, and resist the opposite urge to ban it (which simply pushes usage onto ungoverned personal accounts). A phased rollout gives you value early and control throughout.

| Phase | Goal |
|---|---|
| 1. Pilot | Prove value with a small group on a governed plan and scoped use cases. |
| 2. Expand | Add policy, training, and role-based access as you widen the user base. |
| 3. Govern | Layer in monitoring, data-loss prevention, and a shadow-AI sweep. |
| 4. Optimize | Measure impact, close feedback loops, and retire the tools you replaced. |

## 4.2 Phase 1 — Pilot

- Stand up a **commercial plan** (Team or Enterprise) from day one — never pilot company work on personal accounts.
- Pick a small group of **champions** and two or three concrete, low-sensitivity use cases.
- Set baseline guardrails: SSO if available, default permission behavior, and a clear list of data that is off-limits.
- Define what success looks like — time saved, quality, adoption — so the expand decision is evidence-based.

## 4.3 Phase 2 — Expand

- Publish an **acceptable-use policy** (starter in 4.6) before widening access.
- Turn on **role-based access** and define who can use higher-reach surfaces like Cowork or Code.
- Train users on surface selection and on what never to paste — the human layer is your biggest variable.

- Provide a simple way to request new use cases and connectors so adoption stays inside the guardrails.

## 4.4 Phase 3 — Govern

- Route audit logs to your **SIEM** and review them; capture agentic actions, not just chats.
- Add **data-loss prevention / redaction** for sensitive data classes, especially on agentic and browser surfaces.
- Run a **shadow-AI sweep**: identify personal accounts used for work and migrate that activity onto the commercial plan.
- Govern **connectors and MCP servers** — approve them deliberately; an unvetted connector bypasses your controls.
- Confirm **deprovisioning** works end-to-end via SCIM.

## 4.5 Phase 4 — Optimize

- Measure ROI against your Phase 1 baseline and report it to leadership in plain terms.
- Expand to new use cases that proved out, and **retire** the legacy or shadow tools Claude replaced.
- Re-run your vendor evaluation periodically and after major provider announcements (Part 3).

## 4.6 Acceptable-Use Policy Starter

A workable AI acceptable-use policy does not need to be long. At minimum, make these explicit:

| Section | What to specify |
|---|---|
| Approved access | Which plans and surfaces are sanctioned, and the rule that personal accounts are not for company work. |
| Prohibited data | The data classes that must never be entered (e.g., regulated records, secrets, customer PII) absent specific approval and controls. |
| Human review | Where human review is mandatory before output is used — especially for code, customer-facing content, and anything touching sensitive data. |
| Connectors & tools | How to request and get approval for connectors, MCP servers, and integrations. |
| Incident reporting | What to do if sensitive data is shared by mistake, and who to tell — fast, blameless reporting beats silent exposure. |

| | |
|---|---|
| **BEST PRACTICE** | Pair the policy with a one-page “what not to paste” reminder. Most real exposure comes from well-meaning employees moving fast, not from attackers — make the safe path the obvious one. |

PART 5

# Quick Reference

Cheat sheet, checklist, and a vendor question one-pager

## 5.1 Surface-Selection Cheat Sheet

| Surface | Best for | Note |
|---|---|---|
| Chat (apps) | Q&A, drafting, analysis of pasted content | Lowest reach; the everyday default. |
| Cowork | File/document work for non-developers | Scoped to a folder; finishes whole tasks. |
| Claude Code | Working inside a codebase | Permission controls for shell actions. |
| API / Platform | Building Claude into products/tools | You control data flow and region. |
| Agentic | Acting across live web apps | Highest reach; reserve for vetted, supervised workflows. |

## 5.2 Pre-Deployment Security Checklist

- [ ] Company use is on a **commercial plan** (Team/Enterprise), not personal accounts.
- [ ] **SSO** (SAML/OIDC) enforced; **SCIM** provisioning and deprovisioning verified.
- [ ] Role-based access configured; higher-reach surfaces gated to approved users.
- [ ] Data-handling terms confirmed; **ZDR** in place if regulated data is involved.
- [ ] Compliance needs mapped (SOC 2 / ISO / HIPAA) and verified via the Trust Center.
- [ ] Audit logging enabled and exported to your **SIEM**; agentic actions captured.
- [ ] Data-loss prevention / redaction applied for sensitive data classes.
- [ ] Connectors and MCP servers governed by an approval process.
- [ ] Acceptable-use policy published; “what not to paste” guidance distributed.
- [ ] Shadow-AI sweep completed; personal-account usage migrated.
- [ ] Incident-reporting path defined and communicated.

- [ ] Re-assessment scheduled (periodic + after major provider announcements).

## 5.3 Vendor Question One-Pager

- Is our data used to train models? Default or opt-in? Does it vary by plan?
- What is the data-retention period? Is Zero Data Retention available?
- Which certifications are held? Are reports available under NDA?
- Is a HIPAA BAA available if we need one? What does it restrict?
- SSO and SCIM supported? What access roles exist?
- Who are the subprocessors? Can data stay in a specific region?
- What audit logging is exposed? Can it stream to our SIEM?
- How are incidents disclosed, and how are models/versions controlled?
- What do the contract terms say about liability and indemnification?



*Sources: Anthropic Trust Center (trust.anthropic.com) | Anthropic Privacy Center (privacy.claude.com) | Claude Platform & Code documentation (platform.claude.com, code.claude.com) | Claude product pages (claude.com). All product details reflect public documentation as of June 2026 and are subject to change. Izzet Labs LLC is independent and not affiliated with Anthropic.*