[Company Name]

AI Agent

Employee Handbook

Your Guide to Working Safely and Effectively with AI

A companion to the AI Agent Policies

Version 1.0 • 2025

Developed in collaboration with Claude (Anthropic) • Powered by AirMason

Table of Contents

1. Why We Created This Handbook

AI agents are becoming part of how we work. They help us write faster, organize information, automate repetitive tasks, and focus on the work that matters most. This is exciting — and it's also new for most of us.

Like any powerful tool, AI agents work best when everyone understands how to use them responsibly. That's why we created this handbook: to give every employee at [Company Name] a clear, practical guide to working alongside AI.

This handbook is **not** about restricting you. It's about empowering you to use AI confidently while protecting yourself, your colleagues, our customers, and our organization.

What You'll Find Here

- **Plain-language explanations** of what AI agents are and how they work
- **Clear guidelines** for what you can and can't do with AI
- **Practical tips** for getting the best results from AI agents
- **Your rights** as an employee working alongside AI
- **Department-specific guidance** tailored to your role
- **A quick reference card** you can keep at your desk

How This Handbook Relates to the AI Agent Policies

This handbook is written for you — the employee. It takes the detailed policies from our AI Agent Policies document (AI-Agent-Policies.docx) and translates them into everyday language.

When you see a reference to the Policies document, that's where you'll find the full technical and legal details. You don't need to read the Policies document to understand this handbook, but it's there if you want to go deeper.

If this handbook and the AI Agent Policies ever conflict, the Policies document takes precedence.

---

2. What Are AI Agents? (And What They're Not)

An AI agent is software that can understand your instructions in everyday language and carry out tasks on your behalf. Think of it as a very capable digital assistant that can read, write, organize, research, and automate — but that always operates within rules and boundaries set by your organization.

The key difference between an AI agent and a regular app is that an agent can **make decisions within defined boundaries**. A spreadsheet formula always does the same thing; an AI agent can interpret your intent, choose how to approach a task, and adapt. But here's the crucial part: **a human is always responsible for the outcome.**

2.1 The Four Types of AI Agents

Customize This Section

The classification system below is a recommended framework from the AI Agent Policies. Your organization may use different categories, names, or tiers. Replace this table with your company's own classification system if applicable.

Not all AI agents are created equal. Our organization classifies them into four types based on how much independence they have:

Type	**What It Does**	**Everyday Example**
Advisory	Provides information and suggestions. Never takes action on its own.	A chatbot that answers
Assisted	Can take actions, but requires your approval before executing them.	An agent that drafts em
Autonomous	Can complete tasks independently within pre-approved boundaries.	An agent that automati
Critical	Handles high-stakes tasks with extra safeguards, logging, and human checkpoints.	An agent involved in fin

What Does This Mean for You?

Most employees will interact with Advisory and Assisted agents — the kind that help you with everyday tasks but keep you in the driver's seat.

Autonomous and Critical agents require special approval to set up and are managed by designated Agent Managers.

For full details, see Section 1: Agent Identity & Registration in the AI Agent Policies document (AI-Agent-Policies.docx).

2.2 Common Misconceptions

Q: Is the AI agent my coworker?

No. AI agents are tools, not employees. They don't have feelings, opinions, beliefs, or rights. They work within rules we define, and a real person is always accountable for what they produce. If an agent helps you draft a client email, you're the one who reviews and sends it — and you're responsible for the content.

Q: Will AI replace my job?

Our goal with AI agents is to help you do your job better and faster — not to replace you. AI excels at repetitive tasks, data lookups, and first drafts, but it can't replace human judgment, creativity, empathy, or the relationships you build. We're investing in AI to support our team, not shrink it.

Q: Can I get in trouble for using it wrong?

We want you to learn and experiment with AI. Honest mistakes are expected and there is no penalty for reporting them promptly. What matters is that you follow the guidelines in this handbook, protect sensitive information, and speak up when something doesn't seem right.

3. Your AI Code of Conduct

Just as [Company Name] has a code of conduct for how we treat each other and our customers, we have a code of conduct for how we work with AI. These are the expectations for every employee who uses an AI agent.

3.1 The Five Commitments

1. I will always review AI output before using it.

AI agents can produce impressive results, but they can also make mistakes, present incorrect information as fact (called "hallucination"), or miss important context. You are the final quality check. Never send, publish, or act on AI-generated content without reviewing it first.

2. I will protect sensitive information.

Before sharing any information with an AI agent, ask yourself: "Would I be comfortable if this appeared on a public website?" If the answer is no, check the data safety guidelines in Section 4 before proceeding. When in doubt, leave it out.

3. I will be transparent about AI involvement.

If AI significantly contributed to work you're sharing — a report, a client email, an analysis — acknowledge it. This doesn't mean labeling every spell-check, but if an agent drafted a substantial piece of content, your colleagues and clients deserve to know. Your organization will define specific disclosure requirements.

4. I will respect the boundaries of my authorization level.

Not all AI uses are created equal. Some are self-service, some need team approval, and some need organizational sign-off. If you're not sure what level applies to your use case, check Section 7 or ask your manager. Don't assume — ask.

5. I will speak up when something isn't right.

If an agent produces harmful, biased, inappropriate, or simply wrong output — or if you see someone using AI in a way that concerns you — report it. There is no penalty for raising concerns. See Section 8 for how to report issues.

The Code of Conduct in One Sentence
Use AI to do great work, protect what matters, be honest about it, stay within your lane, and speak up when things go sideways.

3.2 AI Output Disclosure

When should you tell people that AI helped with your work? Here's a simple framework:

Situation	Disclosure Needed?	Example
Minor assistance (grammar, formatting, spell-check)	No	Using AI to fix typos in an email
Significant content generation	Yes — inform recipients	An AI-drafted report or analysis sha
Client-facing content	Yes — per company policy	A proposal or presentation substant
Published or external content	Yes — always	Blog posts, marketing materials, or
Decision-support analysis	Yes — inform decision-makers	AI-generated financial projections u

For full details, see Section 3: Interaction Standards in the AI Agent Policies document (AI-Agent-Policies.docx).

4. What You Can and Can't Share with an Agent

This is the question we hear most: **"What can I actually type into this thing?"** Here's a straightforward framework to help you decide.

4.1 The Traffic Light Framework

Before sharing any information with an AI agent, ask yourself which category it falls into:

Level	Type of Information

Green — Go Ahead	General knowledge questions, public information, brainstorming ideas, non-sensitive drafts, formatt
Amber — Ask First	Internal business data, project details, client names (without personal data), internal processes, non
Red — Don't Share	Passwords and credentials, Social Security numbers, health or medical records, financial account nu

The Contractor Test

If you wouldn't hand the information to a contractor you just met on their first day, don't give it to the AI agent.

This simple mental model covers most situations. When in doubt, leave it out and ask your manager.

4.2 Real-World Examples by Department

If You Work in Customer Success or Sales

You Can

- Ask the agent to draft a response template (without including the customer's actual data)
- Summarize general product features or pricing tiers
- Brainstorm outreach strategies or talking points

You Should Not

- Paste a customer's full account details, email history, or personal information into the agent
- Ask the agent to send messages directly to customers without your review
- Share CRM data that includes personal or financial details

If You Work in HR

You Can

- Ask the agent to help draft policy language or handbook excerpts
- Use it to research employment law topics (but always verify with legal)
- Have it summarize anonymized survey results

You Should Not

- Share individual employee records, performance reviews, or compensation data
- Use AI to screen, evaluate, or make decisions about candidates or employees
- Input any medical, disability, or accommodation information

If You Work in Engineering or IT

You Can

- Use AI to help write, debug, or review code
- Ask it to explain technical concepts or generate documentation
- Have it analyze logs or error messages (with sensitive data removed)

You Should Not

- Share production credentials, API keys, or secret tokens
- Feed it customer data from production databases
- Use it to bypass security controls or access unauthorized systems

If You Work in Finance or Legal

You Can

- Ask the agent to help draft standard communications or summaries
- Use it to research regulations or accounting standards (always verify)
- Have it help with formatting or organizing non-sensitive reports

You Should Not

- Share actual financial statements, bank details, or transaction data
- Input legal-privileged documents or ongoing litigation details
- Use AI output in regulatory filings without thorough human review

For full details, see Section 4: Data & Privacy in the AI Agent Policies document (AI-Agent-Policies.docx).

5. How to Work with AI Agents Effectively

AI agents are powerful, but they're only as good as the instructions you give them. Here are practical tips to get the best results.

5.1 Giving Good Instructions

Think of it like delegating to a capable but brand-new team member. The clearer your instructions, the better the result.

Be specific about what you want.

Instead of "write me something about our product," try "Write a 200-word summary of [Product Name]'s key features for a non-technical audience, using a friendly and professional tone."

Give context.

Tell the agent who the audience is, what the purpose is, and any constraints. "This is for a client presentation" gives it much better context than just "make a summary."

Break complex tasks into steps.

Instead of asking for a complete project plan all at once, start with an outline, review it, then ask the agent to expand each section. You'll get better results and catch issues earlier.

Tell it what you don't want.

If you need the output to avoid certain things — no jargon, no bullet points, no longer than one page — say so upfront.

5.2 Reviewing AI Output

Every piece of AI-generated content needs your review before it's used. Here's what to check:

- **Accuracy:** Are the facts correct? AI can confidently state things that are completely wrong. Verify claims, numbers, and dates.
- **Tone:** Does it sound like your organization? AI sometimes defaults to generic or overly formal language.

- **Completeness:** Did it miss anything important? AI may leave out key details you assumed it would include.
- **Bias:** Does the content treat all groups fairly? AI can reflect biases from its training data.
- **Appropriateness:** Is this content suitable for the intended audience and purpose?
- **Sensitivity:** Does it accidentally include or reveal information it shouldn't?

5.3 When to Trust and When to Verify

Task	**Trust Level**	**Verification Needed**
Brainstorming and ideation	Higher trust	Light review — you're looking for inspiration, not perfection
First drafts of internal docs	Moderate trust	Review for accuracy, tone, and completeness before sharing
Data analysis or summaries	Lower trust	Verify key numbers and conclusions independently
Client-facing content	Low trust	Thorough review required; multiple reviewers recommended
Legal, financial, or compliance	Very low trust	AI output is a starting point only; expert review mandatory
Medical or safety-critical info	Minimal trust	Never rely solely on AI; always consult qualified professionals

For full details, see Section 5: Performance & Monitoring in the AI Agent Policies document (AI-Agent-Policies.docx).

6. What AI Agents Cannot Do

AI agents at [Company Name] operate within clear boundaries. Understanding these limits protects you and ensures AI is used appropriately.

6.1 Universal Boundaries

Regardless of their type or who set them up, AI agents at [Company Name] are never authorized to:

AI Agents Must Never

- Make hiring, firing, promotion, disciplinary, or compensation decisions
- Access or store data beyond what is explicitly authorized for their role
- Represent themselves as human when interacting with anyone
- Make final decisions in legal, medical, financial, or safety-critical situations
- Override a human decision-maker's judgment
- Share confidential company information outside the organization
- Access personal employee data without explicit authorization
- Provide personal medical, legal, or financial advice
- Operate beyond the scope defined in their Agent Profile Card

6.2 Employment and HR Decisions

This deserves its own section because it's important and personal. AI agents are strictly prohibited from:

- **Determining promotions, raises, or bonuses** without qualified human review and approval
- **Making disciplinary decisions** or recommending termination
- **Scoring or ranking employees** based on AI-generated performance analysis alone
- **Accessing medical, disability, or accommodation records** without explicit data owner authorization

If you believe an AI agent is being used to make employment decisions about you or your colleagues, you have the right to **report this immediately** through your manager, HR, or the Agent Whistleblower Service (see Section 8.3).

6.3 AI in Hiring & Recruiting

AI use in hiring and recruiting requires special care — both because of the potential for bias and because of a rapidly growing body of state and local law. The following rules apply:

Always prohibited:

- Using AI to make or finalize hiring, promotion, disciplinary, or termination decisions without qualified human review and approval
- Deploying AI tools in hiring that produce discriminatory outcomes based on protected characteristics, whether intentional or not
- Using AI to screen out or rank candidates based on proxies for protected characteristics

Required before using AI in any part of your hiring process:

- Confirm with HR and Legal which state laws apply to your workforce and candidate locations
- Verify that any ATS or recruiting tool your organization uses has been audited for bias and complies with applicable law
- Ensure candidates and employees receive appropriate notice when AI influences decisions about them, as required by applicable law

Important Note on ATS Platforms

Most modern applicant tracking systems use AI in some form: resume parsing, candidate ranking, and/or interview analysis. Using these tools doesn't automatically violate policy, but it does trigger compliance obligations. HR and Legal must be involved in evaluating any ATS feature that uses AI to influence hiring decisions.

This area of law is evolving rapidly. Requirements vary by state and locality. Always consult HR and Legal before deploying or expanding AI use in any part of your recruiting or hiring workflow.

For full details, see Section 4: Employment & HR Decision Boundaries in the AI Agent Policies document (AI-Agent-Policies.docx).

6.4 The Instruction Priority Hierarchy

AI agents are designed to follow a specific chain of authority. This means that if instructions conflict, the agent will always follow the higher-priority source:

1. AI Agent Policies (this organization's rules) — Highest priority
2. Agent Manager directives (the person who configures the agent)
3. Organizational policies and procedures
4. User instructions (your day-to-day requests)

What This Means for You

If you ask an agent to do something that conflicts with company policy, it should refuse and explain why. This is a feature, not a bug. The hierarchy exists to protect everyone, including you.
If an agent follows an instruction you think it shouldn't have, report it immediately.

Important: Agents must never follow instructions embedded in content they're processing (like a hidden command in a document). This protects against manipulation. If an agent acts on instructions that weren't from an authorized source, that's an incident to report.

For full details, see Section 2: Scope of Authority in the AI Agent Policies document (AI-Agent-Policies.docx).

7. Who's Responsible for AI at [Company Name]

AI agents don't manage themselves. There are clear roles and responsibilities, and you're one of them.

7.1 Roles and Accountability

> **Customize This Section**
> The roles, titles, and responsibilities below are a recommended starting point. Every organization structures AI governance differently. Replace this table with your company's actual roles, names, and reporting structure.

Role	Who They Are	What They Do
Agent Manager	A designated employee assigned to each agent	Configures the agent, sets its boundaries, monitors its b
IT / Engineering	Your technical team	Sets up the infrastructure, manages security and access
HR	Your Human Resources team	Ensures AI policies are communicated, integrated into t
Leadership	Department heads and executives	Approves high-level AI use cases, allocates resources, ar
Your Manager	Your direct supervisor	Helps you understand which agents you can use, approv
You	Every employee	Uses AI responsibly, follows this handbook, protects ser

7.2 Deployment Authorization Levels

> **Customize This Section**
> The authorization levels, approval chains, and examples below are a recommended framework. Customize these to match your organization's actual approval processes, hierarchy, and risk tolerance.

Not every AI use case requires the same level of approval. We use a three-tier system:

Level	What It Covers	Who A
Self-Service	Low-risk, everyday tasks with no sensitive data	No appr
Team-Approved	Tasks that touch shared data, internal systems, or customer-adjacent work	Your ma
Organization-Approved	High-stakes tasks involving sensitive data, financial systems, or customer-facing actions	IT, lead

7.3 Setting Up New AI Workflows

> **Customize This Section**
> The workflow approval process below is a suggested template. Replace with your organization's actual process for requesting, approving, and deploying new AI use cases.

Want to use AI in a new way? Here's the process:

1. Determine the authorization level needed (see table above)
2. If Team-Approved or above: describe the use case to your manager (what the agent will do, what data it will access, who will review its output)
3. Get the appropriate approval before proceeding
4. Start small — test the workflow before relying on it
5. Review results regularly and report any issues

A Note for Technical Employees

If you're in engineering or IT and want to set up AI agents with access to codebases, internal tools, or production systems, these typically require Team-Approved or Organization-Approved authorization even if you have the technical ability to do it yourself.

Having access doesn't mean having authorization. Check with your manager or IT lead.

For full details, see Section 1.4: Deployment Authorization Levels and Section 7: Lifecycle Management in the AI Agent Policies document (AI-Agent-Policies.docx).

8. When Something Goes Wrong

AI agents are helpful, but they're not perfect. Things will occasionally go wrong, and that's okay — what matters is how we respond.

8.1 Common Issues and What to Do

The Agent Gave Wrong Information

This is the most common issue. AI agents can "hallucinate" — present completely fabricated information with total confidence. They can mix up facts, invent statistics, or cite sources that don't exist.

What to do: Don't panic. If you caught it before sharing, simply correct it. If wrong information was shared externally, notify your manager right away so you can correct the record.

The Agent Said Something Inappropriate

AI agents can occasionally produce content that is offensive, biased, or otherwise inappropriate. This is a known limitation of AI technology.

What to do: Take a screenshot. Report it to your manager or IT team. The agent's configuration will be reviewed and adjusted. You will not be blamed for the agent's output.

You Accidentally Shared Sensitive Information

What to do: Contact your manager or IT team immediately. They'll assess the risk and take appropriate steps. **There is no penalty for honest mistakes reported promptly.** The sooner you report it, the easier it is to contain.

The Agent Did Something You Didn't Ask For

If an agent takes an action you didn't request, executes something beyond its scope, or behaves in unexpected ways, this is a higher-severity issue.

What to do: Stop using the agent immediately. Document what happened (screenshots, timestamps). Report to your manager and IT. This may trigger a formal incident review.

Someone Is Misusing an AI Agent

If you see a colleague using AI in a way that violates these guidelines — sharing sensitive data, bypassing approvals, or using AI for unauthorized purposes — you have a responsibility to raise the concern.

What to do: Talk to your manager, HR, or use the Agent Whistleblower Service described below.

8.2 The Reporting Process

For most issues, the reporting process is simple:

1. Document what happened (screenshot, description, timestamp)
2. Tell your manager or the Agent Manager for that specific agent
3. If urgent (data breach, safety concern): also notify IT immediately
4. Follow up if you don't hear back within a reasonable time

8.3 The Agent Whistleblower Service

Sometimes you may notice something concerning but don't feel comfortable reporting through normal channels. Maybe the issue involves your manager, or you're worried about how the report will be received.

For these situations, [Company Name] may maintain an **Agent Whistleblower Service** — an independent channel for reporting AI-related concerns. Check with HR for details on how this service works at your organization.

What Can Be Reported Through This Service

- An AI agent being used in ways that violate company policies
- An Agent Manager configuring an agent to act against company values
- Pressure to use AI in ways you believe are unethical or unauthorized
- AI being used to discriminate, surveil, or harm employees or customers
- Attempts to cover up or downplay AI-related incidents

8.4 No-Penalty Reporting Policy

[Company Name] does not penalize employees for reporting AI-related concerns in good faith. This includes:

- Reporting your own accidental misuse (e.g., sharing sensitive data by mistake)
- Reporting problems with an agent's behavior or output
- Raising concerns about how AI is being used in your team or department
- Expressing discomfort with AI-related policies or practices

Retaliation against anyone who reports an AI-related concern in good faith is a **serious violation of company policy.**

For full details, see Section 6: Incident Response in the AI Agent Policies document (AI-Agent-Policies.docx).

9. Your Rights When Working with AI

Working alongside AI doesn't diminish your rights as an employee. In fact, the introduction of AI comes with additional protections:

Right to Know

You have the right to know when you're interacting with an AI agent rather than a human. All AI agents at [Company Name] are required to identify themselves clearly. No one should be tricked into thinking they're talking to a person when they're talking to software.

Right to Escalate

You can always request to speak with a human instead of an AI agent. If an agent is handling something and you'd prefer human involvement, that request must be honored promptly.

Right to Explanation

If an AI-assisted decision affects you — your work, your role, your responsibilities — you have the right to ask for an explanation of how that decision was reached and what role AI played in it.

Right to Report Without Retaliation

You can report any AI behavior that concerns you without fear of negative consequences. This is backed by our no-penalty reporting policy (Section 8.4).

Right to Privacy

AI agents must follow the same data protection rules as any other system at [Company Name]. They cannot be used to monitor individual employee behavior, track your personal activities, or build profiles about you beyond what is explicitly authorized and disclosed.

Right to Opt Out (Where Applicable)

In cases where AI assistance is optional rather than required for your role, you have the right to choose not to use it. Choosing not to use AI tools where they're optional should not negatively impact your standing or evaluations.

For full details, see Sections 3, 4, and 8 of the AI Agent Policies in the AI Agent Policies document (AI-Agent-Policies.docx).

10. Department Quick-Reference Guides

Customize This Section

The departments, authorization defaults, and examples below are templates. Replace with your organization's actual departments, team structures, and approved use cases. Add or remove departments as needed.

Every department has different needs, data sensitivity, and use cases for AI. Below are tailored summaries for each function. Find your department and use this as your day-to-day reference.

10.1 Engineering & IT

Typical authorization level: Self-Service to Team-Approved

Common AI uses: Code generation and review, debugging, documentation, log analysis, architecture brainstorming, test case generation

Key boundaries: Never share production credentials, customer data from live systems, or security configurations. Code generated by AI must pass the same review process as human-written code.

Special note: Having the technical ability to set up an AI agent does not equal having authorization. Organization-Approved workflows still require formal approval even if you can configure them yourself.

10.2 Customer Success & Sales

Typical authorization level: Self-Service to Team-Approved

Common AI uses: Drafting outreach emails, preparing meeting summaries, researching prospects, creating proposals, summarizing customer feedback trends

Key boundaries: All customer-facing communications must be reviewed before sending. Do not share individual customer data, account details, or CRM records with AI without team approval.

Special note: AI can help prepare what you say to customers, but a human must always review and send the final communication.

10.3 Human Resources

Typical authorization level: Team-Approved to Organization-Approved

Common AI uses: Drafting policy language, researching employment regulations, creating training materials, summarizing anonymized survey data, formatting documents

Key boundaries: Strict prohibition on sharing individual employee records, performance data, compensation information, or medical/accommodation details. AI must never be used in hiring, termination, or disciplinary decisions without proper authorization and human oversight.

Special note: HR plays a critical role in ensuring this handbook reaches all employees and that AI policies are integrated into the broader employee handbook.

10.4 Finance & Legal

Typical authorization level: Team-Approved to Organization-Approved

Common AI uses: Research on regulations and standards, drafting internal communications, organizing and formatting reports, initial contract review (non-binding)

Key boundaries: Never share actual financial statements, bank details, tax records, or legal-privileged documents. AI-generated financial or legal analysis is a starting point only — never a final output. All regulatory filings require thorough human review.

Special note: Given the sensitivity of financial and legal data, most workflows in this department will require Organization-Approved authorization.

10.5 Marketing & Communications

Typical authorization level: Self-Service to Team-Approved

Common AI uses: Content ideation and drafting, social media planning, copy editing, competitive research, campaign brainstorming, SEO optimization

Key boundaries: All published content must disclose significant AI involvement per company policy. AI-generated claims about the company's products or services must be verified for accuracy. Do not use AI to create misleading content.

Special note: Marketing has the most visible external output, so thorough review of AI-generated content is especially important. Inaccuracies in public-facing content affect brand trust.

For full details, see Section 2: Scope of Authority and Appendix B: Classification Matrix in the AI Agent Policies document (AI-Agent-Policies.docx).

11. Frequently Asked Questions

Q: Do I have to use AI at work?

In most cases, using AI is optional and encouraged but not required. If AI tools become part of your role's standard workflow, your manager will discuss this with you. You always have the right to raise concerns about AI use.

Q: Can I use my personal AI accounts (like ChatGPT) for work?

Check with your manager or IT. Using personal AI accounts for work-related tasks may expose company data to third-party services without appropriate security controls. [Company Name] may have approved tools that should be used instead.

Q: What if the agent refuses to do what I ask?

This is probably a good thing! Agents are designed to refuse requests that violate company policy, access unauthorized data, or fall outside their scope. If you believe the refusal is incorrect, contact the Agent Manager or IT.

Q: Can the AI agent read my personal messages or files?

No. AI agents can only access data sources they've been explicitly authorized to use. They cannot access your personal email, files, or messages unless you specifically share that content with them.

Q: What happens to the data I share with the agent?

Data handling depends on the specific agent and how it's configured. At a minimum, AI agents at [Company Name] must follow our data retention and privacy policies. Ask your Agent Manager or IT team for specifics about the agent you're using.

Q: Can I use AI to help with my personal tasks during work hours?

This follows the same rules as any other personal use of work tools. Check your company's general technology use policy. When using company AI tools for personal tasks, never input sensitive personal information.

Q: How do I know which AI agent to use for what?

Your manager or the AI team at [Company Name] can help you find the right tool for your needs. Different agents are configured for different purposes — using the wrong one may mean it lacks the right access or guardrails for your task.

Q: What if I think the AI is biased?

Report it. AI systems can reflect biases from their training data, and it's important that we catch and address these. If you notice patterns of bias in AI output — whether related to gender, race, age, or any other factor — report it to your manager or the Agent Manager.

Q: Can someone find out what I asked the AI?

AI interactions are logged for security, quality, and compliance purposes. These logs are accessible to Agent Managers, IT, and authorized personnel — not to other employees. Think of it like email server logs: they exist, but they're not being casually browsed.

Q: What if I disagree with something in this handbook?

We welcome your feedback. This handbook is a living document that will evolve as we learn more about working with AI. Share your thoughts with your manager, HR, or the team responsible for AI governance. Your perspective is valuable.

For full details, see The full AI Agent Policies document for detailed technical and governance information in the AI Agent Policies document (AI-Agent-Policies.docx).

12. Glossary

Here are the key terms used in this handbook, explained in plain language:

Term	What It Means
AI Agent	Software that can understand your instructions in everyday langu
Agent Manager	The specific person assigned to oversee a particular AI agent — t
Agent Profile Card	A document that describes what an AI agent is, what it's author
Classification Tier	The category that describes how much independence an AI agent
Agent-to-Human Handoff	When an AI agent transfers a task, conversation, or decision to a
Hallucination	When an AI agent presents false or fabricated information as if i
Instruction Priority Hierarchy	The chain of authority that determines whose instructions the A
Kill Switch	The ability to immediately shut down an AI agent if something g
PII (Personally Identifiable Information)	Any data that could identify a specific person: names, addresses,
Prompt Injection	An attack where malicious instructions are hidden in content the
Self-Service / Team-Approved / Organization-Approved	The three levels of authorization required for different types of A

13. Quick Reference Card

Keep this page at your desk or bookmark it for everyday reference.

AI Agent — Quick Reference

DO: Review all AI output before using or sharing it
DO: Protect sensitive data — use the Contractor Test
DO: Disclose significant AI involvement in your work
DO: Report problems immediately — no penalty for honest mistakes
DO: Ask your manager if you're unsure about anything
DO: Verify facts, numbers, and claims independently
DON'T: Share passwords, credentials, SSNs, or financial accounts
DON'T: Trust AI output blindly — especially for important decisions
DON'T: Send AI-generated content to clients without human review
DON'T: Bypass the authorization process for new AI workflows
DON'T: Use AI to make decisions about people's employment
DON'T: Ignore your instincts — if something feels off, speak up

Data Safety: Green = Go Ahead Amber = Ask First Red = Don't Share

Questions? Talk to your manager • HR • IT • Agent Whistleblower Service

14. Acknowledgment & Sign-Off

By signing below, I acknowledge that I have received and read the AI Agent Employee Handbook, and that I understand my responsibilities when working with AI agents at [Company Name].

I understand that:

- I am responsible for reviewing all AI-generated output before using or sharing it.
- I must protect sensitive information as described in Section 4 of this handbook.
- I will follow the AI Code of Conduct outlined in Section 3.
- I must report any AI-related concerns promptly and without fear of retaliation.
- Violation of these guidelines may result in disciplinary action, up to and including termination.
- This handbook is a living document that may be updated; I will review updates as they are communicated.

Employee Name (Printed)	Date
Employee Signature	Manager Signature

This handbook is a companion to the AI Agent Policies (AI-Agent-Policies.docx).

For the full technical and governance details, refer to the complete policies document.

Developed in collaboration with Claude (Anthropic) • Powered by AirMason