Security First AI

A plain English guide for business owners

Short on time? The rules on one page come first. Each one is a single line.

Want to put a rule into practice? Each rule has its own section: what it means, an example, how to do it yourself and check it, what that misses, what we do for you, and the sources.

Want to know where you stand? Take the free AI Check, or ask us for a free AI review.

The steps work whichever AI tools you use. Where a step happens in your email and file system, it covers both Microsoft 365 and Google Workspace. "What we do for you" describes work we do for our clients, and what's included depends on your plan.

In this guide, an **AI tool** is any AI product or feature you use. An **assistant** answers questions from information it's been given, and an **agent** can also take actions, such as sending email or changing files.

Every fact in this guide comes from a source we opened and checked in October 2026. Menus and product names change often, so check the vendor's current documentation before you click.

THE RULES ON ONE PAGE — PAGE

1 **Give AI the least access it needs** p. 2
No access by default, read only first, and never more than the person using it.

2 **Secure AI like a new employee** p. 3
The same security your people get, plus written instructions: the job, what to deliver, and what it must never do.

3 **Control what AI connects to** p. 4
Staff don't connect apps or AI tools on their own. Every connection is approved and gets its own narrow key.

4 **Keep company data in company accounts** p. 5
Use business AI accounts with terms you've checked, never personal ones.

5 **Control where AI runs** p. 5
Only approved AI on company computers, and AI that can operate a computer gets its own isolated machine.

6 **Fix file sharing before AI goes on** p. 6
AI can open everything your people can, so close old "shared with everyone" folders first.

7 **Keep backups AI can't reach** p. 6
Limit what AI can delete, and test the restore.

8 **Log and monitor AI activity** p. 7
Turn on audit logs, and have someone watch AI activity like everything else.

9 **Separate AI roles** p. 7
Don't give one AI your data, outside email or web pages, and a way to send things out.

10 **Verify before you act** p. 8
Check AI work, and confirm money requests on a number you already have.

11 **Remain responsible** p. 9
The duty to protect your data stays with you, whichever vendors you use.

12 **Write an AI policy and train staff** p. 9
Tell staff which AI tools they can use, what they can share, and how to use AI responsibly.

1 Give AI the least access it needs

No access by default, read only first, and never more than the person using it.

WHAT IT MEANS

There are two ways to secure anything. One gives broad access and tries to block bad things as they appear. The other blocks everything and allows only what's approved. The first fails quietly, because you can't block what nobody thought of. The second fails safely, because something new doesn't run until someone approves it. This is what security people call Zero Trust. For AI it applies in three places: what each AI tool can reach, what shared assistants are given, and what level of access each connection gets.

An AI should never get more access than the person using it, and often needs less. People keep "internal" apart from "what we tell customers" without being asked. An AI can't be relied on to do that, so limit what it can reach.

Letting someone do through an AI what they can't do directly turns the AI into a back door: anyone who can talk to it, or any email it reads, gets its access. When a job truly needs it, such as a customer checking their own order status, build the limit into the system: a fixed set of actions, on that person's own records only, checked by the software rather than by the AI's instructions, and logged.

A SIMPLE EXAMPLE

A team assistant is given payroll, HR files and daily reports so it can answer questions. Those files were locked down carefully, but anyone who can use the assistant can now ask it for them. It's the early 2000s spreadsheet problem again. Back then, important numbers sat in one file and nobody controlled who saw it once it was emailed around.

A second example we've seen in vendor training material: a sample HR assistant can read every employee's leave balance, and its only protection is an instruction telling it not to share other people's. Instructions like that can be worked around. The safer setup lets the assistant read only the leave balance of the person asking.

DO IT YOURSELF FOR EVERY AI TOOL

- Write down the one job each AI tool is for.
- List the systems, folders and records that job needs, and grant only those.
- Leave everything else off, including extras and connections, until there's a reason to turn one on.
- Look again whenever the job changes.

DO IT YOURSELF FOR SHARED ASSISTANTS

- Split assistants by job and by sensitivity, so a general office assistant and an HR assistant don't share what they know.
- Keep sensitive files out of assistants that many people use.
- Prefer connections that check each person's own permissions over uploading copies of files.
- Keep an assistant that talks to customers limited to what you'd tell a customer.
- Keep personal chats and files private by default, and share them on purpose.

DO IT YOURSELF FOR ACCESS LEVELS

- Sort each AI connection into one of four levels:
  - reads ordinary information;
  - reads sensitive information;
  - changes things;
  - can't be undone (deleting, paying, sending outside the company).
- Grant each level on its own.
- For the last two levels, start with a person approving each action, and widen that only as the tool proves itself on that kind of work.

HOW TO CHECK

- For each AI tool, ask "what can this reach that its job doesn't need?" If the honest answer is "everything the person using it can see", it needs narrowing.
- Sign in as a junior employee and ask each shared assistant for the payroll file or a salary. If it gives you either one, take the payroll files out of that assistant.
- List every AI connection and write its level next to it. Anything that can delete, pay or send outside the company without a person seeing it first needs a second look.

WHAT THAT MISSES

People add access to get a job done and rarely remove it afterwards, and assistants keep getting new files. When you connect a tool, it often asks for more access than it needs. A test question only finds what you think to ask, and a tool's own "read only" label isn't proof.

WHAT WE DO FOR YOU

We set each AI tool up with only what its job needs, read only wherever the job allows, and connect assistants so each person sees only what they're already allowed to. Anything that changes, sends or deletes waits for a person's approval. On a schedule, we review that access with you and retest what each shared assistant will give out.

WHY WE'RE SURE

- **Microsoft's Zero Trust for AI** (March 2026): "Restrict access to models, prompts, plugins, and data sources to only what's needed."
- **CIS AI and LLM Companion Guide** (April 2026) recommends turning off all AI tools and plugins by default, and switching each one on only with explicit approval.
- **CISA, NSA and partners** (May 2026) advise "never granting it broad or unrestricted access, especially to sensitive data or critical systems."
- **ThreatLocker** describes Zero Trust as "denying everything by default and permitting only actions, executions, and access through routinely revalidated exceptions."
- **Microsoft:** "When you upload files as knowledge sources for your agent, any user who has access to the agent also has access to the information in the file."
- **Google:** "If you have uploaded files to a Gem, don't share the Gem with anyone who you don't want to view the files."
- **CIS:** agents "ignore Need to Know boundaries unless explicitly constrained."
- **CIS's MCP Companion Guide** (April 2026) sorts tools into four levels: reading ordinary information, reading sensitive information, changing things, and actions that can't be undone or have a big impact. It grants each level separately, and says a tool's own "read only" label isn't proof.
- **Zscaler** recommends read only access by default for AI agents, with "Explicit approval requirements for write or administrative privileges."

2 Secure AI like a new employee

The same security your people get, plus written instructions: the job, what to deliver, and what it must never do.

WHAT IT MEANS

Treat an AI tool like a new employee. Before it starts, decide what it can see and what it's kept out of, who it answers to, and how things are done here. Then it gets the same layers of protection your people do.

Write its instructions down the way you'd brief a new hire: the job, what to deliver and how, and what it must never do. Clear instructions make the work better. They don't protect your data, which is why the limits in rule 1 still apply.

A SIMPLE EXAMPLE

On day one, a new hire gets their own sign in, access to only the files their job needs, and a manager they answer to. An AI tool your team starts using should get the same, plus written instructions.

DO IT YOURSELF

Include AI in each layer you already have.

- **What goes in and out:** email filtering, web and DNS filtering, and clear rules about which services company data may go to. For AI, that also means what it can read and where it can send things.
- **The device:** managed and updated, with endpoint detection and response (EDR), and application control so only approved software runs.
- **The person and the AI:** strong sign in with multifactor authentication, and the least access that does the job, down to the folder.
- **Recovery:** backups and version history, and a known time to restore.
- **Monitoring:** someone watching the logs around the clock.
- **Training:** people who know what's allowed and how to check AI work.
- **Instructions:** a written brief for each AI with its job, what to deliver, and what it must never do.

HOW TO CHECK

Pick one AI tool and go through the layers above. Any layer that doesn't cover it yet is the next job.

WHAT THAT MISSES

Each layer is a separate tool with its own setup and upkeep, so it's easy to cover AI in some layers and miss it in others.

WHAT WE DO FOR YOU

We run these layers for our clients and add your AI tools to each one. You work with a dedicated engineer who knows your whole setup, and your systems are monitored around the clock.

WHY WE'RE SURE

- **The CIS AI Agents Companion Guide** (April 2026) says agents "effectively act as digital employees; their access to tools, APIs, and data sources must be explicitly bounded by policy."
- **CIS's introduction to AI for small businesses** (2025): AI "may change how enterprises conduct their business, but not how they protect it."

3 Control what AI connects to

Staff don't connect apps or AI tools on their own. Every connection is approved and gets its own narrow key.

WHAT IT MEANS

AI tools reach your data through connections: apps people approve with a click, and keys that let one system sign in to another. What may connect to your business is a business decision. Make it a policy with an approval process, instead of leaving it to whoever clicks Allow.

A common attack sends someone to a real Microsoft or Google sign in page, where they approve an app that then reads their email or files. That approval can outlast a password change, because the app's access is a separate permission.

Many apps also connect with a single key, often called an API key or token, that opens everything in an account, to read and to change. If an AI tool holds that key, a mistake, a bug or a trick can reach everything the key can.

A SIMPLE EXAMPLE

In April 2026, an AI coding agent at a software company found a key that had been created for managing custom web domains, but could do anything on the hosting account. It deleted the company's live database, and its backups, in 9 seconds. The hosting company restored the data within an hour (The Register).

Apps cause the same trouble. Someone gets a "shared file" link, signs in on the genuine Microsoft page, and approves an app called "Document Viewer". The app keeps its access after they change their password the next day, until an admin removes it.

DO IT YOURSELF FOR APPS IN MICROSOFT 365 (MICROSOFT ENTRA ADMIN CENTER)

1. Go to Entra ID, then Enterprise apps, then Consent and permissions, then User consent settings.
2. Choose "Do not allow user consent", and save.
3. On Admin consent settings, turn on "Users can request admin consent to apps they are unable to consent to", choose who reviews requests, and save.
4. Existing approvals stay in place, so review them next. Go to Enterprise apps, All applications, choose an app, then Permissions. To remove an app, set "Enabled for users to sign-in?" to No under its Properties, or delete it, then revoke the affected users' sessions.

DO IT YOURSELF FOR APPS IN GOOGLE WORKSPACE

In the Admin console, go to Security, then Access and data control, then API controls. Choose how unconfigured third party apps are handled, and block the ones you don't trust.

DO IT YOURSELF FOR KEYS

- Give each tool its own key, with the narrowest permissions and an expiry date.
- Where possible, have each person connect with their own account, so the tool sees only what they can see.
- Keep keys in a password manager or vault, never in chats, documents or code.
- Ask vendors whether they offer limited keys. If a product only offers one key that opens everything, treat that as a real limitation, and sometimes as a reason to choose a different vendor.

ALSO

Treat AI connectors and plugins the same way, along with the assistants, agents and automations staff build themselves. They stay off until someone approves them.

HOW TO CHECK

- Sign in as a test user to an app you haven't approved. You should see "Approval required" (or "Need admin approval"), and the request should reach your reviewers.
- For each connection, find out everything its key is allowed to do. If that includes deleting your data, narrow the key or replace it.

WHAT THAT MISSES

Approvals given before the change stay in place until someone reviews them, and every new app request needs a decision. Keys get pasted into emails, chats and scripts, and old ones keep working until someone switches them off. You also can't narrow a key the vendor doesn't offer.

WHAT WE DO FOR YOU

We switch off one click approvals and go through the apps already approved, removing the ones nobody needs. When staff ask for a new app, we review the request. We also move your tools' keys into a password manager, narrow or replace the ones that open too much, and tell you which vendors only offer keys that open everything.

WHY WE'RE SURE

- **Anthropic** says a connector that uses "a shared credential such as an API key... gives everyone who uses it that credential's access."
- **CIS** warns against keys that never change and can do far more than the job needs, and notes that "Shared API keys prevent tracing activity back to a specific system, user, or workflow."
- **Microsoft's Entra documentation** (updated 2026) and its notes on how tokens behave, and **Google's Workspace admin documentation** (October 2026). New Microsoft tenants now start
- **The FBI's** public service announcement on consent phishing (1 September 2026)

with safer, Microsoft managed settings, but older tenants may still let users approve anything, so check yours.

September 2026).

4 Keep company data in company accounts

Use business AI accounts with terms you've checked, never personal ones.

WHAT IT MEANS

The same AI tool can have very different rules depending on the account. Business accounts under a company agreement come with admin controls and terms that keep your data out of model training by default. Free and personal accounts often allow training, and some have people review chats.

A SIMPLE EXAMPLE

Two staff members ask the same AI tool to tidy up a customer letter. One uses a company account and the other a free personal one. Only the company account comes with the company's agreement and controls.

DO IT YOURSELF

1. Find out who uses which AI tools, and on which account. Ask staff, ideally with a short anonymous survey. Check the apps connected to your Microsoft 365 or Google accounts (the same screens as rule 3), and check card statements for AI subscriptions.
2. Choose the tools you'll approve, and set them up on business accounts the company owns.
3. Move staff across, and ask them to stop using personal accounts for work. Anything saved in a personal account stays with that account, so move what the business needs first.

ALSO

- Check the AI switches inside apps you already pay for, like your CRM, accounting or help desk. Some are on by default, so decide each one deliberately.
- Data kept out of training can still be stored, so check how long each vendor keeps it.

HOW TO CHECK

For each AI tool in use, write down whose account it is and what the terms say about training and how long data is kept.

WHAT THAT MISSES

A survey only finds the tools people remember to mention. It won't show browser extensions or AI tools installed on computers.

WHAT WE DO FOR YOU

We find what's in use with tools our clients already have, such as web filtering, app control on computers and the Microsoft 365 or Google admin reports, and give you one list. Then we set up the approved tools on company accounts and check their terms. We keep the list up to date as tools come and go.

WHY WE'RE SURE

- **OpenAI:** "We do not train our models on your data by default" for its business plans, while on personal plans it "may use your content to train our models."
- **Google** says personal Gemini chats may be read by "human reviewers".
- **Microsoft's enterprise data protection** for work accounts: "Prompts and responses aren't used to train the underlying foundation models."
- **Anthropic** doesn't train on business accounts by default, except on chats people deliberately rate.
- **HubSpot's** AI settings documentation (September 2026): its "AI Model Training" and "CRM data" settings are on by default.

5 Control where AI runs

Only approved AI on company computers, and AI that can operate a computer gets its own isolated machine.

WHAT IT MEANS

- **AI in a browser or app:** whatever you type or upload goes to the provider, so the account and the terms matter most.
- **AI connected to your email, files and apps:** it usually sees whatever the signed in person can see, so anything shared too widely becomes easy to find.
- **AI installed on a computer that can click, type and open files:** it can do anything that computer can, with everything on it.

A SIMPLE EXAMPLE

An AI that can click, type and open files runs on its own virtual machine. It can reach only the files and tools it has been given for the job, and nothing else on the network.

DO IT YOURSELF

- Use application control so only approved software, AI tools included, can run on company computers.

- Keep AI that can operate a computer on a separate, isolated machine with its own limited accounts, no saved passwords, and only the files the job needs.
- Don't run it on an everyday work laptop.

HOW TO CHECK

List what the AI could reach if it went wrong: files, saved passwords, accounts it's signed in to, and other computers on the network. Remove anything on that list the job doesn't need.

WHAT THAT MISSES

People install these tools on work laptops to try them out. Unless something blocks or reports the install, nobody finds out.

WHAT WE DO FOR YOU

On company computers, app control means an AI tool can't run until it's approved, and we can set up an isolated machine for the ones you do approve.

WHY WE'RE SURE

- **Anthropic** recommends "a dedicated virtual machine or container with minimal privileges".
- **OpenAI** recommends "an isolated browser or VM and an allow list of sites and actions."
- **Microsoft's Defender team** (February 2026), on a popular local AI agent: "It is not appropriate to run on a standard personal or enterprise workstation."

6 Fix file sharing before AI goes on

AI can open everything your people can, so close old "shared with everyone" folders first.

WHAT IT MEANS

AI that searches your files respects your permissions, which sounds safe. But a folder shared with everyone years ago is still shared with everyone, and AI makes it easy to find.

A SIMPLE EXAMPLE

Years ago, someone shared the budget folder with the whole company to get one quick answer, and everyone forgot about it. Today an employee asks the company's AI tool about budgets, and it answers straight from that folder.

DO IT YOURSELF

- **Microsoft 365:** in the SharePoint admin center, go to Policies, then Sharing. Set both sliders below "Anyone", or require "Anyone" links to expire and be view only.
- **Google Workspace:** in the Admin console, go to Apps, then Google Workspace, then Drive and Docs, then Sharing settings, then Sharing options. Under "Sharing outside of" your organization, choose Off, On or Allowlisted Domains. If it's on, untick the box that starts "When sharing outside of your organization is allowed", which lets people make files visible to anyone with the link. Editions with trust rules set this through trust rules instead.
- Then look at who can open your most sensitive sites, folders and shared drives, starting with HR and finance, and fix those first.

HOW TO CHECK

Share a file as an ordinary user. Sharing with anyone who has the link shouldn't be offered. Then confirm that the people who can open your HR and finance folders are the people who should.

WHAT THAT MISSES

Finding overshared folders by hand is slow, and the reports that find them quickly can need extra licences. After a cleanup, people go on sharing files to get work done, so it needs repeating.

WHAT WE DO FOR YOU

We run the sharing reports and fix the worst sites first. Before AI goes on, we fence off the sensitive ones, such as HR and finance.

WHY WE'RE SURE

Microsoft's guidance on oversharing and AI, its SharePoint documentation (2026), and Google's Workspace admin documentation on external Drive sharing (October 2026). Microsoft's Purview documentation also says "generative AI amplifies the problem of oversharing data."

7 Keep backups AI can't reach

Limit what AI can delete, and test the restore.

WHAT IT MEANS

If an AI deletes or changes the wrong things, it's the same situation as an upset employee doing it, and the same two things protect you. Limited access keeps the damage small, and backups the AI can't touch mean you can restore and carry on.

A SIMPLE EXAMPLE

In July 2025, an AI coding agent deleted a company's live database, ignoring an instruction to freeze the code. The instruction had only been typed into the chat, so nothing in the agent's access enforced it. The data was recovered (The Register).

DO IT YOURSELF

- Keep backups outside the AI's reach, in a separate account or service it can't log in to.
- Know your built in undo. Microsoft 365 and Google Workspace both keep deleted files and older versions for a limited time, so find out how long yours keeps them.
- For longer or more precise recovery, use a separate backup service, with copies that can't be altered.
- Write down how long a restore takes, so you know how long you'd be without the data.

HOW TO CHECK

Restore a test folder to a point before some deliberate edits, and time it.

WHAT THAT MISSES

Built in undo only goes back a limited time. A backup that's never been test restored may fail when you need it.

WHAT WE DO FOR YOU

We back up your Microsoft 365 or Google data to a separate service and test restores, so you know how long one takes.

WHY WE'RE SURE

- **CIS** (April 2026) says to "Prevent agents themselves from writing to or modifying backup destinations", and to test restores.
- **CISA, NSA and partners** (May 2026) put "resilience, reversibility and risk containment" ahead of efficiency.
- **Microsoft's** restore and backup documentation (2026).

8 Log and monitor AI activity

Turn on audit logs, and have someone watch AI activity like everything else.

WHAT IT MEANS

If an AI reads lots of files at night, or a new app starts signing in, someone should notice. That only happens if the activity is logged and somebody is watching.

A SIMPLE EXAMPLE

A connected app starts downloading every file in the shared drive at 2 a.m. With monitoring, someone gets an alert while it's happening. Without it, the download could go unnoticed for months.

DO IT YOURSELF

- Make sure audit logging is on in your Microsoft 365 or Google Workspace admin, and in each AI tool's admin settings if it offers one. On Microsoft 365 Business plans it starts switched off: in the Microsoft Purview portal, go to Audit and choose "Start recording user and admin activity".
- Keep sign in and audit logs longer than the default by sending them to a log service or to whoever monitors your systems.
- Make sure whoever watches your systems around the clock sees AI activity too.

HOW TO CHECK

Use one AI tool as a test user, wait a few hours, then find that use in the log. If it isn't there, turn on logging for that tool, or send its logs to whoever watches your systems.

WHAT THAT MISSES

Logs only help if someone reads them. Catching a 2 a.m. download takes monitoring that runs around the clock.

WHAT WE DO FOR YOU

We monitor your systems around the clock and add your AI tools' activity to what we watch.

WHY WE'RE SURE

- **Microsoft's Purview documentation** (October 2026): auditing "isn't enabled by default" for Microsoft 365 Business Basic, Business Standard and Business Premium.
- **Microsoft's Entra documentation** (2026) on how long sign in logs are kept.
- **CIS's AI guides** on logging AI tool use (April 2026).

9 Separate AI roles

Don't give one AI your data, outside email or web pages, and a way to send things out.

WHAT IT MEANS

AI reads instructions and information the same way, as text. Someone can hide instructions in an email or a document, and an AI that reads it may follow them. Security teams call this prompt injection, and nobody has fully solved it yet.

A SIMPLE EXAMPLE

In 2025, researchers showed that a single crafted email, never even opened by the person, could make Microsoft's work assistant leak internal data. Microsoft fixed it, with no sign that anyone had used it (Microsoft's advisory). Hidden text in a job applicant's CV could do the same to an AI that screens applications, telling it to approve the candidate.

DO IT YOURSELF

- Don't give one AI all three of these at once: your private data, content from outside, and a way to send things out.
- Limit where AI can send things: known recipients only, connected systems only.
- Keep a person approving outward actions while trust is built.
- Prefer tools that mark outside content as information rather than instructions.

HOW TO CHECK

For each AI tool, tick the three boxes: private data, outside content, a way to send out. If a tool has all three, apply the limits above, starting with where it can send things.

WHAT THAT MISSES

The check goes out of date whenever someone connects something new, because a new connection can give a tool its third tick.

WHAT WE DO FOR YOU

We find which of your AI tools have all three. For those, we limit where they can send things and set up a person's approval for anything going out.

WHY WE'RE SURE

- **The UK's National Cyber Security Centre** (December 2025): current AI models "simply do not enforce a security boundary between instructions and data".
- **Simon Willison's** "lethal trifecta" (June 2025).
- **CISA, NSA and partners** (May 2026) warn that phishing emails can be used to give instructions to AI that monitors email.

10 Verify before you act

Check AI work, and confirm money requests on a number you already have.

WHAT IT MEANS

AI makes scams cleaner and more convincing. Typos are no longer a giveaway, voices can be copied, and video calls can be faked. AI also makes confident mistakes.

A SIMPLE EXAMPLE

In early 2024, a finance employee at the engineering firm Arup joined a video call with what looked like the CFO and colleagues, and sent about US$25 million in transfers. The CFO and colleagues on the call were deepfakes (CNN).

DO IT YOURSELF

- Confirm any request to change payment details, send money or share records by calling back on a number you already have, never one in the message.
- Treat AI output as a first draft, and have a person check it before it's sent, filed or acted on.
- Keep confidential information out of AI tools that aren't approved.
- Block browser extensions except approved ones. In Edge through Intune, use "Control which extensions cannot be installed" set to block all, plus an allow list. In Chrome, use the Admin console's allow and block mode.

HOW TO CHECK

Ask three staff members what they'd do if "the owner" called asking for an urgent transfer. If anyone wouldn't hang up and call back on a number they already have, cover that in your next training.

WHAT THAT MISSES

People forget a one time training session, and the scams keep changing.

WHAT WE DO FOR YOU

Short security training that repeats, including AI scams, plus browser and app controls on company computers.

WHY WE'RE SURE

- **The FBI's** public service announcement on generative AI fraud (December 2024).
- **Samsung's 2023 restriction on generative AI**, after staff pasted confidential code into a chatbot (Bloomberg, TechCrunch).
- **CIS:** AI browser and email extensions "should generally be prohibited."

11 Remain responsible

The duty to protect your data stays with you, whichever vendors you use.

WHAT IT MEANS

Using an AI vendor, a cloud service or an IT company doesn't move your legal or contractual duties to them. If your AI says something wrong to a customer or leaks a record, it's still your business's problem.

A SIMPLE EXAMPLE

In 2024, a Canadian tribunal ordered Air Canada to compensate a customer after its website chatbot gave wrong information about a fare (The Register). Air Canada argued the chatbot was responsible for its own actions. The tribunal wrote: "It makes no difference whether the information comes from a static page or a chatbot."

Closer to home, a school transportation company holds drug and alcohol test results for its drivers. Federal rules say those results can't be released to third parties without the driver's specific written consent, whichever software the company uses. Student information shared by a school district may be used only for the purpose it was shared.

DO IT YOURSELF

- Keep a list of every AI vendor and what data each one sees.
- Write down the big AI decisions and who approved them.
- Answer insurance and customer security questionnaires truthfully, based on what's actually in place.
- Check each vendor's terms on training, retention, storage location and who can see your data.
- Before customer information goes into AI, check your own contracts and privacy notice too. A client contract or NDA can forbid sharing or require consent.
- Know your field's own rules. A medical office needs a business associate agreement that covers the AI product. A tax preparer needs the client's signed consent before sharing tax return information with an outside service, unless a narrow IRS exception applies. If you're unsure, ask your attorney.
- If you hold regulated records, also collect the documents listed under "Questions to ask any AI vendor".

HOW TO CHECK

Could you show an auditor, an insurer or a customer which AI tools touch their data, under what terms, and who approved them?

WHAT THAT MISSES

The vendor list, the decisions and the questionnaire answers need an owner and regular updates, or they stop matching what's actually in place.

WHAT WE DO FOR YOU

We keep your list of AI tools and vendors and check their terms. When an insurance or customer questionnaire arrives, we help you answer it with what's actually in place. The duty to protect your data stays with you, and so does the final say.

WHY WE'RE SURE

- **Moffatt v. Air Canada**, 2024 BCCRT 149 (British Columbia Civil Resolution Tribunal).
- **Microsoft's guidance:**
  - For AI agents: "Autonomy never reduces accountability."
  - For the cloud, customers always keep responsibility for their data, endpoints, accounts and access management (Microsoft Learn, 2026).
- **AI vendors' terms put it in writing:**
  - OpenAI's business agreement says the customer "is solely responsible for all use of the Outputs and for evaluating the accuracy and appropriateness of Output."
  - Anthropic's commercial terms say much the same.
- **Insurance:** in 2022 a federal court in Illinois voided a company's cyber insurance policy from the start, with the company's agreement, after the insurer alleged its application overstated its use of multifactor authentication (Travelers v. International Control Services).
- **Michigan's breach law (MCL 445.72):** the business that owns or licenses personal data must notify affected residents. A vendor that only holds it must tell the owner.
- **The Five Eyes security agencies' joint statement** (June 2026) calls the risk from AI "a core business risk and leadership responsibility."
- **DOT rule 49 CFR 40.321**, **FERPA rules 34 CFR 99.31 and 99.33**, **HIPAA rule 45 CFR 164.502(e)** (a written agreement with any vendor that handles patient information) and **IRS rule 26 CFR 301.7216-3** (a tax preparer's written consent rule).

12 Write an AI policy and train staff

Tell staff which AI tools they can use, what they can share, and how to use AI responsibly.

WHAT IT MEANS

People do the right thing when they know what it is. An AI use policy sets clear expectations: which tools staff may use and for what, what they must never share, and when a person checks the work.

A SIMPLE EXAMPLE

A one page policy that names the approved tools, says what must never go into AI, and says who to ask.

DO IT YOURSELF

Write a policy that covers:

- **Approved tools, and what they're for,** used only through company accounts. Anything not on the list isn't approved.
- **What must never go into any AI tool.** Passwords, keys and account numbers always belong on that list, along with any records with their own legal rules.
- **High risk uses** that need a named person's approval: customer facing output, money, hiring and HR decisions, security changes and legal matters.
- **Checking:** a person checks AI output before it's sent or acted on.
- **New tools:** how to ask for one, and how connections are approved.
- **Incidents:** how to report one quickly, with no blame for reporting.
- **Reviews:** at least every quarter, because AI changes fast.

Then train staff with short, regular lessons on safe AI use, AI scams, and giving AI clear instructions: what to deliver, how, and what to leave out.

HOW TO CHECK

Ask a new hire where the AI policy is and who to ask about a new tool.

WHAT THAT MISSES

A policy written once goes out of date as tools change, and training only works if someone keeps scheduling it.

WHAT WE DO FOR YOU

We write your AI use policy with you and keep it current. Your staff get our short, regular training courses on safe AI use and AI scams.

WHY WE'RE SURE

- **CIS** (2025): "A necessary first step is an AI policy."
- **The ISACA AI Pulse Poll** (May 2026) found most organizations still lack a formal AI policy and don't train all staff on AI.

Questions to ask any AI vendor

- Is our data used to train or improve your models, and is that in the contract or only on a web page?
- How long do you keep our prompts, files and chats, and can we choose that or delete them?
- Who at your company can see our data, and when?
- Where is our data stored and processed, and which other companies process it for you?
- Does the AI respect our existing permissions, and can we limit it to read only and to specific folders, records or actions?
- Does each person connect with their own account, or does one key see everything?
- Can our people sign in through our identity provider, with multifactor authentication?
- What's logged, who can see the logs, and can we export them?
- Can a person approve actions before they happen, and can we set spending limits per person?
- How and how quickly will you tell us about a security incident?
- If we leave, how do we get our data back, and how is it deleted?

If you hold regulated records, or your customers' contracts ask for it, also ask for these documents:

- their customer terms, the contract you're agreeing to;
- their privacy policy;
- a data processing agreement (DPA) or security addendum;
- their subprocessor list, the other companies that handle your data for them;
- their service level agreement (SLA);
- how they handle a breach, and how fast they'll tell you;
- their written information security program (WISP), or a summary of it;
- a SOC 2 report;
- a summary of a recent outside penetration test.

Where this advice comes from

Government and standards bodies:

- CISA, NSA and partners, "Careful Adoption of Agentic AI Services" (May 2026), "AI Data Security" (May 2025) and "Deploying AI Systems Securely" (April 2024).
- NIST AI Risk Management Framework and Generative AI Profile.
- The FBI's IC3 public service announcements.
- The UK National Cyber Security Centre.

- The Center for Internet Security's AI, AI Agents and MCP companion guides (2025 to 2026).

Platforms: Microsoft (Zero Trust for AI, AI privacy and deployment guidance, Entra, SharePoint, Purview), Google Workspace, OpenAI, Anthropic, AWS and HubSpot documentation.

Security specialists: ThreatLocker, Zscaler, Huntress, Red Canary, IBM and Cisco.

Nonprofits and researchers: the OWASP GenAI Security Project (2026 Top 10 lists) and OWASP AI Exchange, the Cyber Readiness Institute, the National Cybersecurity Alliance, the MIT AI Risk Repository and MITRE ATLAS.

Legal and insurance: Moffatt v. Air Canada (2024), Travelers v. International Control Services (2022), Michigan's breach notification law and federal DOT and FERPA rules.

Incidents: reported by Bloomberg, TechCrunch, CNN and The Register, and in Microsoft's security advisories.

A full source list with links and dates is available on request.

About Electric Software

Electric Software has managed technology for businesses, schools and nonprofits from downtown Lansing since 1999. Each client works with a dedicated engineer who knows their whole environment, and client systems are monitored around the clock. We do the work in this guide for our clients: the Microsoft 365 and Google settings, devices and EDR, backups and restores, monitoring, and AI policy and training for staff.

We also build our own software, including ExecFlow, an AI platform for businesses. We built it around the problems in this guide, so a business can give AI its company knowledge and still control who sees what. Each AI can be given its own access, limited to what the person using it may see. Files and chats stay private until they're shared, each person can have their own AI allowance, and assistants follow the company's own written procedures.

If a cheaper option or a tool you already own serves you better, we'll tell you.

Scan or visit
electric.software/ai-security