AI agent policy, one page

Template. Replace the items in [brackets] with your own, then keep only what your company really enforces.

[Company] uses AI agents every day. These are the rules they work under. Each important rule is also an automatic check that stops the action, so it does not depend on anyone remembering it.

What they may do

- Read and sort the office inbox, draft replies, and prepare documents and code.
- Build, test and deploy changes to our own websites and products, then check the live page after every deploy.
- Prepare a task in the browser on [Owner]'s laptop, and stop before the final click that publishes, submits or pays, unless [Owner] approved that exact step.

What they may never do

- Decide money, prices, contracts or new clients. Those stay with [Owner].
- Comment, send messages or connect with people on LinkedIn. [Company]'s own posts go out only after [Owner] approves each one.
- Send an email that has not passed automatic checks: the address must be real and published, must not have bounced before, and the message must not repeat one already sent.
- Show a password or an API key. A guard refuses any command that would print one.
- Post AI-written text on a platform that does not allow it.

Who answers

- [Owner name], [role]: [email address]. Every alert stays open until it is closed with proof of what was checked.

Template by Devaland, free to copy and adapt.
https://devaland.com