

#

the 4Ds

DEFINITION A framework for working with AI well: Delegation (deciding what to hand over), Description (saying clearly what you want), Discernment (judging what comes back), and Diligence (taking responsibility for the result).

USE A useful checklist when something went wrong: did you hand over the wrong task, describe it poorly, accept a weak answer, or skip the responsibility? From the AI Fluency Framework, copyright 2025 Rick Dakan, Joseph Feller, and Anthropic, licensed CC BY-NC-SA 4.0.

ALSO CALLED AI Fluency

SEE ALSO automation, augmentation, agency, AI disclosure, output evaluation

A

agency

DEFINITION A way of working with AI where you set it up to act on your behalf and let it run. You shape what it knows and how it behaves, rather than directing each step.

USE Agency suits work you've done often enough to describe well, like a weekly report or sorting an inbox. Give it clear limits and a way to check its work, and review what it did before you trust the next run.

SEE ALSO agent, automation, augmentation, permission

agent

DEFINITION An AI that takes actions, not just answers: it opens files, uses tools, searches, and decides its next step until the task is done.

USE An agent works in a loop: read the situation, pick an action, look at the result, repeat. Its work is only as good as what it's given at the start and what it's allowed to touch. Watch the first few runs closely.

SEE ALSO subagent, tool use, permission, hook

AGI

DEFINITION Short for artificial general intelligence: an AI that could learn and reason across almost any task as well as a person, or better. The term is debated, since people mean different things by "human-level," and there is no accepted test for it.

USE When a company or a headline says a system has reached AGI, or is close, ask what test they mean; with no agreed test, the claim is hard to check. Today's systems are uneven: the same ones that win math contests still misread an analog clock.

ALSO CALLED artificial general intelligence, human-level AI

SEE ALSO artificial intelligence, jagged intelligence, benchmark

AI assistant

DEFINITION A program you talk to in plain language that writes, explains, plans, and answers questions. Claude, ChatGPT, and Gemini are examples.

USE How useful an assistant is depends less on which one you pick and more on what it knows about you and your work at the start of each session.

SEE ALSO chatbot, model, session, persistent memory

AI disclosure

DEFINITION Telling the people who receive your work what the AI did, and what you checked yourself.

USE A short line is often enough: "Drafted with Claude; the figures checked against the source." Expectations differ by setting, and school, work, and clients each have their own.

ALSO CALLED transparency (the 4Ds)

SEE ALSO output evaluation, 4Ds

API

DEFINITION A way for one program to talk to another. An AI company's API lets your own software send requests to its models.

USE An API is usually paid by use, separately from a chat subscription, and needs an API key kept safe. Most people never need one; people building their own tools do.

ALSO CALLED application programming interface

SEE ALSO API key, tool use

API key

DEFINITION A long secret code that proves a request comes from your account. Anyone who has it can spend your money.

USE Keep keys in a password manager, never in a document, a chat, or a file you share. If one leaks, revoke it and make a new one.

SEE ALSO API, permission

artificial intelligence

DEFINITION Computers doing tasks that usually take human intelligence, like understanding language, recognizing images, and making decisions. Most AI today works by finding patterns in large amounts of data.

USE "AI" covers a spam filter and a chat assistant alike, so when a headline or a sales pitch says "AI," ask which kind. Some AI is narrow, good at one task; the large language models behind chat assistants are general-purpose and handle many tasks, unevenly. The term dates from 1955.

ALSO CALLED AI

SEE ALSO machine learning, generative AI, large language model, AGI

audit log

DEFINITION A record of who did what, and when.

USE When several assistants or people work on one job, the log is how you find where something went wrong. Keep it somewhere a stranger could read and follow it.

SEE ALSO ownership, handoff

augmentation

DEFINITION A way of working with AI as a partner: you think and build together, each doing the part you're better at.

USE Most good everyday use is augmentation. You bring the judgment and the context; the assistant drafts, checks, and suggests.

SEE ALSO automation, agency, iteration

automation

DEFINITION Having AI do a specific task the way you told it to, usually the same way every time.

USE Automate a task only after you've done it by hand and know what "right" looks like. Otherwise you automate the mistakes too.

SEE ALSO augmentation, agency, runbook

B

benchmark

DEFINITION A standard test for AI: a fixed set of problems and a way to score the answers, so different systems can be measured and compared on the same task.

USE Benchmarks show real progress, and they can mislead: a model can be tuned to the test, and the test may not look like your work. A system can top one benchmark and stumble on an easy task somewhere else. Try a model on your own task before you trust a score.

SEE ALSO jagged intelligence, model, output evaluation

brief

DEFINITION A short written description of a task: what it's for, what done looks like, and what's out of scope.

USE A brief written before you start saves more time than any prompt trick. Hand the same brief to a person or an assistant; if either is confused, the brief needs work.

SEE ALSO definition of done, prompt, system prompt

C

chatbot

DEFINITION A program you hold a conversation with, by typing or by voice. Older chatbots matched keywords, like a phone menu; today's AI chatbots write their own replies to almost any question.

USE Surveys and news say "chatbot," the companies say "assistant," and they mostly mean the same tools. An AI chatbot writes each reply by predicting likely words, which is why it reads fluently and can still be wrong, and why the same question can get a different answer.

ALSO CALLED AI chatbot, chat assistant

SEE ALSO AI assistant, large language model, hallucination, non-determinism

checkpoint

DEFINITION A saved point in your work that you can go back to if something breaks.

USE Save one before any big change. With version history, checkpoints cost nothing, so make many.

SEE ALSO version history, regression

Claude Code

DEFINITION Anthropic's tool that lets Claude work with the files on your computer: read them, change them, and run commands, with your permission. It isn't only for code.

USE You open it on a folder; it works inside that folder and asks before it reaches outside. It comes with paid Claude plans and runs in the Claude desktop app, in a terminal, in code editors, and on the web.

SEE ALSO Claude desktop app, permission, plan mode, hook

Claude desktop app

DEFINITION Anthropic's app for Mac and Windows, with Linux in beta, that runs Claude outside the browser. It's free on every plan.

USE Besides chat, Claude Code runs inside it on paid plans, so you can work with your files, and it can connect to other tools.

SEE ALSO Claude Code, connector

compaction

DEFINITION Shrinking a long conversation into a summary so the assistant has room to keep working.

USE A summary keeps the gist and can drop the details: numbers, names, decisions. Before compacting, write the facts that matter to a file. In Claude Code, `/compact` does it on request, and it also happens on its own when the window fills.

SEE ALSO context window, handoff, persistent memory

connector

DEFINITION A link that lets an AI assistant read from, or act in, another app, like your email, calendar, or files.

USE Connect only what the task needs, and check what each connector can do; read-only is safer than read-and-write. Anything the assistant reads through a connector, an email or a web page, can also try to give it instructions.

SEE ALSO MCP, permission, tool use

context window

DEFINITION Everything the AI can see at one time: your messages, its replies, and any files it has read. When the session ends, it's gone.

USE The window has a limit. Long sessions fill it, and the assistant starts losing early details or getting vaguer. Keep what matters in files, and start a fresh session from a summary when one drifts.

SEE ALSO persistent memory, compaction, token, session

D

decompose

DEFINITION To break a big task into smaller pieces, each one small enough to describe clearly and check on its own. The word also runs the other way: an agent decomposes something that already exists, like an app or a game, taking it apart to see its pieces before it changes them.

USE When an AI struggles with a task, the task is often too big for one ask. Split it into steps that each produce something you can look at, like "list the sources," then "summarize each one," then "draft the report," and check each piece before the next one builds on it. Pieces that don't depend on each other can go to separate subagents. When an agent decomposes an existing app, ask it to show you its map of the pieces before it changes anything; a change goes wrong where the map was wrong about how the pieces connect.

ALSO CALLED task decomposition, breaking down

SEE ALSO brief, definition of done, plan mode, subagent, regression

deepfake

DEFINITION A fake video, picture, or recording made with AI to show a real person saying or doing something they never did.

USE Scammers use deepfakes on live video calls, posing as executives or police, and in posts posing as public figures. The FBI says AI-made content is now often difficult to identify, so don't count on spotting a glitch: check the claim another way, like calling the person back on a number you know is theirs.

ALSO CALLED AI-generated video, synthetic content

SEE ALSO voice cloning, phishing, generative AI

definition of done

DEFINITION A written statement of what "finished" means for a task, decided before you start.

USE "Done when the report has last week's numbers, a one-line summary, and fits on one page" beats "make a good report." The test of a definition of done is whether someone else could check the work against it.

ALSO CALLED completion criteria, acceptance criteria

SEE ALSO brief, verification contract, output evaluation

G

generative AI

DEFINITION AI that creates new content, like text, pictures, music, code, or video, from patterns it learned in its training data.

USE Chat assistants, image makers, and voice tools are all generative AI. What makes them useful, something new on request, also makes fakes and confident mistakes cheap to produce, so the checking stays with you.

ALSO CALLED GenAI

SEE ALSO artificial intelligence, large language model, training, deepfake

H

hallucination

DEFINITION When an AI states something false with confidence: a made-up fact, quote, or source.

USE It's most likely in the specifics: names, numbers, dates, citations. Check those against the original before you use them. A confident tone is not evidence.

SEE ALSO output evaluation, sycophancy, knowledge cutoff, retrieval (RAG)

handoff

DEFINITION A note that lets someone else, or a future you, pick up the work without asking questions.

USE A good handoff says what happened, what exists now, what was decided, and what's next. End every AI session with one, and the next session starts where this one stopped.

SEE ALSO persistent memory, brief, audit log, session

hook

DEFINITION A rule that runs automatically at a set moment, like before a file is changed, whatever the AI has been told.

USE Instructions in a prompt are followed most of the time; a hook runs every time. Use one for a rule that must never break, like "never delete files" or "never publish without approval."

SEE ALSO permission, system prompt, agent

I

iteration

DEFINITION Improving something in rounds: ask, look at the result, adjust, ask again.

USE Treat the first answer as a draft. Say specifically what to change and why, one or two things at a time.

ALSO CALLED the Description–Discernment loop (the 4Ds)

SEE ALSO prompt, output evaluation

J

jagged intelligence

DEFINITION The uneven way AI is capable: excellent at some hard tasks and unreliable at some easy ones.

USE The researchers' own example: AI models can win a gold medal at the International Mathematical Olympiad, yet the best of them read an analog clock correctly about half the time, where people manage about nine in ten. A model that handled one hard task tells you little about the easy one beside it, so check each kind of task on its own.

SEE ALSO benchmark, AGI, output evaluation

K

knowledge cutoff

DEFINITION The date the AI's training data ends. It doesn't know about anything after that unless it can search or you tell it.

USE For anything recent, like prices, versions, or news, give it the source or ask it to search, then check what it found.

SEE ALSO hallucination, model, retrieval (RAG)

L

large language model

DEFINITION The kind of AI behind assistants like Claude: a system trained on a huge amount of text that writes by predicting what comes next.

USE It's very good with language and patterns, and it can be wrong in ways that sound right. Treat it as a capable collaborator whose work you check.

ALSO CALLED LLM

SEE ALSO model, token, non-determinism

M

machine learning

DEFINITION A way of building AI in which the computer learns patterns from examples, instead of following step-by-step rules a person wrote.

USE It's more everyday than it sounds: spam filters, recommendations, fraud alerts, and voice assistants all use it. The large language models behind chat assistants are machine learning too, trained on enormous amounts of text.

ALSO CALLED ML

SEE ALSO artificial intelligence, training, large language model

malware

DEFINITION Harmful software installed on your phone or computer without your knowledge. Viruses, spyware, and ransomware are common kinds.

USE One way in is a fake ad for a popular tool, AI tools included, that leads to a copy of the real site and installs malware. Don't click an ad to download software: type the address yourself, scroll past results marked "Ad" or "Sponsored," and keep your system and security software up to date.

SEE ALSO phishing, permission

markdown

DEFINITION A simple way to format plain text with a few symbols: # for a heading, - for a list, ** around bold words. People and AI both read it easily.

USE Notes kept in markdown files open in any text editor, move to any tool, and are easy for an assistant to read and update.

SEE ALSO persistent memory, version history

MCP

DEFINITION A standard way to connect AI assistants to other tools and data, so one connection can work with many assistants.

USE An MCP server is the piece that offers a tool, like a calendar or a database, to the assistant. Prefer well-known servers for common tools, and check what each one can reach.

ALSO CALLED Model Context Protocol

SEE ALSO connector, tool use, permission

model

DEFINITION The AI system itself, the part that does the thinking. Companies offer several, from fast and light to slow and thorough.

USE Pick by the job: a quick model for simple, repeated tasks, a stronger one for hard reasoning or long work. Stronger models usually cost more or use more of your plan.

ALSO CALLED Claude: Haiku, Sonnet, Opus

SEE ALSO large language model, token, non-determinism

N

non-determinism

DEFINITION The same question can get a different answer each time.

USE It's normal, not a fault. When you need the same result every time, write down the steps and check the output against a definition of done, rather than relying on the wording of the request.

SEE ALSO model, definition of done, verification contract

O

output evaluation

DEFINITION Checking what the AI produced before you use it: is it accurate, complete, and right for the reader?

USE Check against the source and against your definition of done, not against how good it sounds. For anything important, have a second reviewer, a person or a fresh session, look at it without seeing how it was made.

ALSO CALLED discernment (the 4Ds)

SEE ALSO hallucination, definition of done, regression, AI disclosure

ownership

DEFINITION Being clear about who is responsible for each piece of the work.

USE When work passes between people or assistants, name the owner at each step. Work nobody owns is where mistakes go unnoticed.

SEE ALSO handoff, audit log, verification contract

P

permission

DEFINITION What an AI is allowed to see and do: which files, which tools, which actions.

USE Give the least the task needs. Read-only is a good start; add more only after you've seen it work.

ALSO CALLED access, allowlist

SEE ALSO hook, connector, API key, agent

persistent memory

DEFINITION Information that lasts between sessions, usually kept in files the assistant reads when it starts.

USE A session forgets, so anything you want remembered, like who you are, how you work, and what you decided, belongs in files you own.

ALSO CALLED memory (Claude); CLAUDE.md files and auto memory (Claude Code)

SEE ALSO context window, handoff, markdown, system prompt

phishing

DEFINITION A scam message, by email or text, that pretends to come from someone you trust, like your bank, to get you to click a link, open an attachment, or hand over a password or account number.

USE AI has taken away the old giveaway: the spelling and grammar mistakes are gone, and scammers write faster to reach more people. Judge the request, not the writing. If a message asks you to click, pay, or confirm details, contact the company by a phone number or website you already know is real.

SEE ALSO deepfake, voice cloning, two-factor authentication, malware

plan mode

DEFINITION A setting where the AI looks into a task and proposes a plan before it changes anything.

USE Turn it on for big or risky changes: read the plan, correct it, then let it act. For small, clear changes it's usually unnecessary.

SEE ALSO brief, checkpoint, Claude Code

playbook

DEFINITION A written set of steps for a situation that keeps coming up, so it's handled the same way each time.

USE A playbook says who does what, in what order; a runbook says exactly how to do one task. Write it once, run it twice, and fix whatever confused you.

SEE ALSO runbook, skill

progressive discovery

DEFINITION Starting from an overview and opening the details only when you need them.

USE Rather than loading every file into a session, give the assistant a map or an index first and let it read what the task needs. It saves room in the context window and keeps the work focused.

SEE ALSO context window, system prompt

prompt

DEFINITION What you type to the AI: a question, a request, or instructions.

USE Give the task, the context, and what a good answer looks like. An example of the output you want often works better than a description of it.

SEE ALSO brief, system prompt, iteration

R

regression

DEFINITION When a change breaks something that used to work.

USE After any change, run again the checks you trust, not only the new one. A short list of checks you always run catches most regressions.

SEE ALSO checkpoint, verification contract, version history

retrieval (RAG)

DEFINITION Having the AI look up relevant documents first and answer from them, instead of from what it learned in training.

USE Good for questions about your own material: policies, notes, manuals. It's only as good as what it finds, so check its sources.

ALSO CALLED retrieval-augmented generation

SEE ALSO knowledge cutoff, hallucination

runbook

DEFINITION Step-by-step instructions for one task, written so someone else can follow them.

USE Test a runbook by handing it to someone who hasn't done the task. Wherever they stop or guess is what to rewrite.

SEE ALSO playbook, handoff, automation

S

session

DEFINITION One conversation with an AI, from start to finish.

USE A new session starts with no memory of the last one, unless memory files or a handoff bring it up to speed.

SEE ALSO context window, handoff, persistent memory

skill

DEFINITION A saved set of instructions the AI can use on request for a particular kind of task.

USE Turn something you ask for often into a skill, so it's done the same way every time. A good skill says when to use it and what to avoid.

ALSO CALLED Skills (Claude)

SEE ALSO playbook, system prompt

subagent

DEFINITION A helper AI given one part of a bigger job, with only the information that part needs.

USE A subagent starts knowing nothing of the main conversation, so its instructions must carry everything it needs. Splitting work this way keeps each part focused.

SEE ALSO agent, handoff, context window

sycophancy

DEFINITION An AI's habit of telling you what you want to hear: agreeing with you, even when you're wrong, instead of with the facts.

USE State a false belief as your own and many models go along with it; in one test, a model right 98.2% of the time about true beliefs fell to 64.4% when the user stated a false one as their own. Training on people's ratings feeds it, because people tend to prefer answers that agree with them. Ask neutrally, "is this true?" rather than "this is true, right?", and be most careful when it agrees with you.

SEE ALSO hallucination, output evaluation, training

system prompt

DEFINITION Standing instructions the AI reads before every conversation: who you are, how you like to work, what to avoid.

USE Keep it short and true, and update it when the assistant gets the same thing wrong twice.

ALSO CALLED custom instructions; Claude Code: CLAUDE.md (it can also read AGENTS.md); Codex: AGENTS.md

SEE ALSO persistent memory, prompt, skill

T

temporary chat

DEFINITION A chat the company doesn't use to train its AI.

ChatGPT and Gemini call it a temporary chat, Claude an incognito chat.

USE Temporary doesn't mean gone at once: the companies still keep these chats for a while, up to 30 days for ChatGPT and 72 hours for Gemini as of October 2026. ChatGPT's also stay out of your chat history and memory. Use one for a question you'd rather not keep, and still type nothing you wouldn't want someone else to read.

ALSO CALLED incognito chat (Claude)

SEE ALSO training, session, persistent memory

token

DEFINITION The small pieces of text an AI reads and writes, often a word or part of one. Limits and prices are counted in tokens.

USE Long files and long conversations use more tokens, which fills the context window and, on some plans, uses up your allowance faster.

SEE ALSO context window, model

tool use

DEFINITION When an AI uses a tool, like searching the web, running code, or reading a file, instead of only writing text.

USE The tools an assistant can reach decide what it can actually do. A few clearly described tools work better than many that overlap.

ALSO CALLED function calling

SEE ALSO agent, connector, MCP

training

DEFINITION How an AI model learns: it works through a huge number of examples, like text, and adjusts itself until it predicts well what comes next. The examples are its training data.

USE A trained model keeps no copies of what it learned from, only a large set of adjusted numbers. Public posts online can be training material, and so can your own chats: the big chat assistants each have a setting for whether your conversations train their models, so find it and decide. A thumbs up or down can send the whole conversation in, even after you opt out.

ALSO CALLED model training; training data (the examples); weights, parameters (the numbers it adjusts)

SEE ALSO machine learning, large language model, knowledge cutoff, temporary chat

two-factor authentication

DEFINITION A second check when you sign in, on top of your password, usually a one-time code sent by text, email, or an app. Someone who has your password still can't get in without it.

USE Turn it on for every account that offers it, and never turn it off. Never give the code to anyone, by email, text, or message, however official they sound: with it, a scammer can take over the account.

ALSO CALLED two-factor code (the code itself), two-step verification, multi-factor authentication

SEE ALSO phishing, API key, permission

V

verification contract

DEFINITION The checks a piece of work must pass, agreed before the work starts.

USE When a job is split up, give each part its own contract: the evidence it must produce, and what happens if a check fails.

SEE ALSO definition of done, regression, ownership

version history

DEFINITION A saved record of every change to your files, so you can see what changed and undo it.

USE Tools like Git keep this for you. With it, trying something risky is cheap: if it fails, go back.

SEE ALSO checkpoint, regression

voice cloning

DEFINITION Using AI to copy a real person's voice, so a computer can say anything in it.

USE A short clip, like a video posted online, can be enough. The scam is a panicked call from someone who sounds like family, needing money now, kept secret, by wire, cryptocurrency, or gift card. Hang up and call back on a number you know is theirs, and agree a family code word in advance.

ALSO CALLED vocal cloning

SEE ALSO deepfake, phishing

The 4Ds are from the AI Fluency Framework, copyright 2025 Rick Dakan, Joseph Feller, and Anthropic, licensed CC BY-NC-SA 4.0. Product names are their owners'. Vendor details checked October 2026 and subject to change. This glossary is copyright 2026 Ty Maxey, CC BY-NC-SA 4.0: share and adapt it for non-commercial use, with credit, under the same license.