FLASHFEED STUDY GUIDE

AI for Everyone

Study guide by FlashFeed

November 2026 edition · pack version 0.1.19

flashfeedlearningapp.com

AI for Everyone: study guide

A practical, tool-agnostic course for using generative AI well: better prompts, examples, agents, media, documents, research, data, code, verification, privacy, and reusable workflows.

Edition	November 2026 (2026-11-21)
Pack version	0.1.19
Contents	14 lessons · 495 cards in the app · 73 self-test questions · 23 pictures
In the app	flashfeedlearningapp.com/packs/ai-for-everyone

How to use this guide

Take one lesson at a time: read the study guide, then go over any key facts. Where a lesson has a self-test, try it from memory before you check the answer key at the back. Follow the Go deeper links when you want more.

Free, with no ads and no account to make.

Contents

01	A Good Prompt Has Parts	4
02	Learn by Showing	10
03	Roles, Reasons, and Control	16
04	From Chatbots to Agents	22
05	Make with Multiple Media	28
06	Writing and Documents	34
07	Research, Learning, and Decisions	40
08	Data, Spreadsheets, and Code	46
09	Verify and Protect	51
10	Build Your Personal AI Workflow	57
11	A Map of AI	62
12	Words You'll See Everywhere	67
13	What AI Can—and Cannot—Do	72
14	Where Smart People Disagree	76
	Answer key	81
	Sources	86
	Image credits	88
	Useful links	89
	About FlashFeed	90

LESSON 1 OF 14

A Good Prompt Has Parts

Name the Task · Give Relevant Context · Set Constraints and Format · Define Success

WHAT YOU'LL BE ABLE TO DO

- Build a prompt with a clear task, relevant context, useful constraints, and a testable output.

Name the Task

A useful prompt begins with a verb and an outcome. “Help with my trip” leaves the system to guess whether you want ideas, an itinerary, a budget, or packing advice. “Draft a two-day walking itinerary” names an observable deliverable. The wording does not need to be formal; it needs to make the job visible.

Before sending, underline the action and circle the deliverable. If either is missing, add it. For an unfamiliar task, ask the system to restate the assignment and list open questions before it answers. That small check exposes mismatched assumptions while they are still cheap to fix.

Three ideas to retain:

- **Start with an Action:** A strong prompt states what the system should do: draft, compare, classify, extract, critique, plan, or explain. A verb makes the requested behavior easier to evaluate.
- **Name the Deliverable:** The same topic can produce a memo, table, checklist, lesson, or set of questions. Naming the deliverable prevents the system from choosing the container for you.
- **Separate Goal from Topic:** “Retirement” is a topic. “Explain three retirement-account differences to a new employee” is a goal. A goal combines subject, action, and intended result.

Key term: **Task verb** means “the action the system should perform.”

Give Relevant Context

Context is information that would change a competent answer. Audience, purpose, setting, prior decisions, and source material often matter. A beginner needs different explanations than a specialist; a private brainstorming note needs a different tone than a public announcement.

More context is not automatically better. Long, unfiltered dumps can bury the decisive detail, and models do not use every position in a long context equally well. Provide the smallest relevant packet, label its sections, and distinguish facts from preferences. If something is unknown, say so instead of inviting the system to invent it.

Three ideas to retain:

- **Context Changes the Answer:** Audience, purpose, constraints, and source material determine what a useful answer looks like. The same task can need different content in a classroom, clinic, or kitchen.

- **Relevant Beats Exhaustive:** Context should earn its place by changing the output. A focused brief is usually easier to use and verify than a long dump of unrelated history.
- **Label Inputs Clearly:** Headings such as BACKGROUND, SOURCE TEXT, and REQUIREMENTS help separate evidence from instructions and make complex prompts easier to inspect.

Key term: **Audience context** means "who will use or read the output."

Set Constraints and Format

Constraints turn taste into instructions. Length, tone, reading level, required sections, excluded content, deadline, and output format are all legitimate boundaries. The best constraints reflect the real use of the result: a table for comparison, bullets for scanning, or JSON for software.

Constraints can conflict. "Comprehensive," "one paragraph," and "include every exception" may not fit together. Rank the requirements or state which can flex. Prefer positive directions—what to include and how to organize it—while using exclusions for genuine hazards or scope boundaries.

Three ideas to retain:

- **Format Is Part of the Task:** A correct answer can still be unusable if it arrives in the wrong structure. Specify a table, numbered plan, email, schema, or other container when the container matters.
- **Constraints Need Priorities:** Length, completeness, tone, and detail can pull in different directions. Ranking constraints tells the system what should win when all cannot be satisfied.
- **Positive Directions Are Clearer:** "Use three short sections with concrete examples" provides a target. A long list of prohibitions may define only what the output should not be.

Key term: **Hard constraint** means "a boundary the output must satisfy."

Define Success

Acceptance criteria describe what must be true of a satisfactory answer. They might require every recommendation to cite supplied evidence, every recipe to avoid an allergen, or every row to contain a date. Criteria make revision less personal: instead of "make it better," you can point to the missed requirement.

A miniature example often communicates style faster than adjectives. Show one short input and the kind of output you want, then explain which features should carry over. End by asking the model to check its draft against the criteria and flag any item it could not meet. That check is evidence, not a guarantee.

Three ideas to retain:

- **Criteria Make Quality Testable:** Acceptance criteria translate "good" into observable conditions such as required evidence, fields, length, coverage, or prohibited content.
- **Examples Carry Tacit Structure:** A small example can reveal naming, granularity, ordering, and tone that would take many abstract instructions to describe.

- **Self-Checks Still Need Review:** A model can compare its answer with a checklist, but it can also miss its own error. Treat the check as a useful pass, not independent verification.

Key term: **Acceptance criterion** means "a condition the result must meet."

Sources

- Google Cloud — Prompt design strategies

Key terms

Task specification
A concise statement of the action, subject, and intended result.

Context
Information supplied with a request that helps determine a relevant response.

Output format
The structure or representation in which a response should be delivered.

Rubric
A set of criteria used to judge the quality of an output.

Key facts



AI-generated (gpt-image-1.5)

Name the Deliverable

The same topic can produce a memo, table, checklist, lesson, or set of questions. Naming the deliverable prevents the system from choosing the container for you.

> **Separate Goal from Topic**
>
> "Retirement" is a topic. "Explain three retirement-account differences to a new employee" is a goal. A goal combines subject, action, and intended result.



AI-generated (gpt-image-1.5)

Relevant Beats Exhaustive

Context should earn its place by changing the output. A focused brief is usually easier to use and verify than a long dump of unrelated history.

Label Inputs Clearly

Headings such as BACKGROUND, SOURCE TEXT, and REQUIREMENTS help separate evidence from instructions and make complex prompts easier to inspect.

Constraints Need Priorities

Length, completeness, tone, and detail can pull in different directions. Ranking constraints tells the system what should win when all cannot be satisfied.

Positive Directions Are Clearer

“Use three short sections with concrete examples” provides a target. A long list of prohibitions may define only what the output should not be.

Examples Carry Tacit Structure

A small example can reveal naming, granularity, ordering, and tone that would take many abstract instructions to describe.



AI-generated (gpt-image-1.5)

Self-Checks Still Need Review

A model can compare its answer with a checklist, but it can also miss its own error. Treat the check as a useful pass, not independent verification.

Matched pairs

Each item on the left goes with the one on the right.

Task verb → The action the system should perform

Deliverable → The concrete output you expect

Audience context → Who will use or read the output

Situational context → The circumstances that shape the task

Hard constraint → A boundary the output must satisfy

Soft preference → A desirable feature that may flex

Acceptance criterion → A condition the result must meet

Reference example → A sample that demonstrates desired features

Self-test

Answer each question from memory, then check your work against the answer key.

1. Which opening is most actionable?

- A Tell me about plans
- B What do you think?
- C Compare these three plans
- D Plans are interesting

2. What is missing from "Help with my résumé"?

- A A creative persona
- B A specific outcome
- C A model nickname
- D A longer greeting

3. Which detail most changes a safety handout?

- A The font is your favorite
- B The prompt says please
- C The model has a logo
- D The readers are children

4. What is the best context rule?

- A Include details that change the answer
- B Hide uncertainty
- C Repeat the task five times
- D Paste every related document

5. Which instruction defines a format?

- A Be intelligent
- B Think carefully
- C Use your knowledge
- D Return a three-column table

Go deeper

WATCH Start With a Verb and an Outcome 0:30 · YouTube
"Help with my trip." That prompt leaves the system guessing. Ideas? An itinerary? A budget? Packing advice?

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/a-good-prompt-has-parts

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 2 OF 14

Learn by Showing

Zero-Shot Requests · One-Shot Examples · Few-Shot Boundaries · Iterate with Evidence

> WHAT YOU'LL BE ABLE TO DO
>
> - Choose zero-, one-, or few-shot prompting and improve examples through deliberate iteration.

Zero-Shot Requests

A zero-shot request gives instructions without a worked example. It is a sensible starting point for familiar tasks, quick exploration, and cases where the desired format is easy to describe. Zero-shot does not mean context-free: you can still provide source text, constraints, and a rubric.

Use the first result as a probe. If the model understands the task and the cost of an error is low, keep the simple prompt. If formatting drifts or a category is subtle, add an example instead of piling on vague adjectives. The smallest prompt that reliably works is easier to maintain.

Three ideas to retain:

- **Zero-Shot Means No Demonstration:** A zero-shot prompt asks the model to perform a task without showing a completed input-output example. Instructions and context can still be detailed.
- **Simple Tasks Often Start Directly:** For a familiar transformation such as turning notes into bullets, a clear instruction may be enough to discover whether examples are needed.
- **The First Output Is a Probe:** A low-risk first attempt reveals ambiguities in the task, missing context, and format drift. Its failures are information for the next prompt.

Key term: **Zero-shot prompt** means “instructions without a worked example.”

One-Shot Examples

One-shot prompting adds a single demonstration. It is especially useful when the desired output has an unusual shape, label scheme, or tone. Choose an example that reveals the rule rather than an easy case that would be obvious without it.

Separate the example from the real input with headings or delimiters. Keep its content different enough that the model must transfer the pattern rather than copy the answer. After the output, inspect whether the demonstrated feature—not accidental wording—was what carried over.

Three ideas to retain:

- **One Example Can Define a Pattern:** A worked input-output pair can demonstrate the expected labels, level of detail, or transformation more precisely than style adjectives.
- **Examples Need Boundaries:** Clear markers around EXAMPLE INPUT, EXAMPLE OUTPUT, and YOUR INPUT reduce confusion between the demonstration and the live task.

- **Choose a Representative Case:** An example should exercise the rule you care about. An unusually easy or exceptional case can teach the wrong boundary.

Key term: **One-shot prompt** means "instructions plus one demonstration."

Few-Shot Boundaries

Few-shot prompting supplies several demonstrations. Its greatest value is not repetition; it is boundary coverage. Include ordinary cases, a difficult edge case, and—when useful—a near miss. The set can communicate which differences matter and which should be ignored.

Examples consume attention and context, so curate rather than accumulate. Research shows that demonstrations can help through format, label space, and input distribution even when their labels are not perfectly informative; that does not make correctness optional. Accurate, relevant, diverse examples remain the practical default.

Three ideas to retain:

- **Variation Teaches Boundaries:** Several examples can show a category's center, edge, and near miss. Diversity is more informative than repeating nearly identical easy cases.
- **Examples Communicate More Than Labels:** Demonstrations can reveal output format, available labels, and the distribution of inputs, not only a direct reasoning rule.
- **Curate the Set:** More examples use more context and can introduce contradictions. Keep examples relevant, accurate, and distinct enough to add information.

Key term: **Few-shot prompt** means "instructions plus several demonstrations."

Iterate with Evidence

Prompting is closer to testing than spell-casting. Save the original task, collect a few representative inputs, and compare versions against the same rubric. Changing one major variable at a time makes it easier to learn why a result improved or regressed.

When an answer fails, diagnose the failure class: misunderstood task, missing context, wrong format, weak example, unsupported claim, or unsuitable tool. Rewrite the smallest responsible piece. A reusable prompt is not a frozen incantation; it is a versioned instruction that earns trust on known cases.

Three ideas to retain:

- **Prompting Is Test-Driven:** A prompt version is useful when it performs well on representative cases and a stated rubric, not merely when one output feels impressive.
- **Change One Major Variable:** Holding the test cases steady while changing one instruction or example makes the effect easier to interpret.
- **Diagnose Before Rewriting:** Different failures need different fixes. Missing evidence is not repaired by a more enthusiastic tone, and a wrong format may not need a new model.

Key term: **Prompt regression** means "a previously good case fails after a change."

Sources

- Min et al. — Rethinking the Role of Demonstrations
- Wang et al. — In-Context Learning from Invalid Demonstrations
- Google Cloud — Prompt design strategies

Key terms

Zero-shot prompting
Requesting a task without supplying a completed demonstration.

Demonstration
A worked example included in a prompt to show how an input should map to an output.

Few-shot prompting
Prompting with a small set of worked demonstrations.

Prompt iteration
A cycle of testing, diagnosing, revising, and comparing instructions.

Key facts

> **Simple Tasks Often Start Directly**
>
> For a familiar transformation such as turning notes into bullets, a clear instruction may be enough to discover whether examples are needed.

> **The First Output Is a Probe**
>
> A low-risk first attempt reveals ambiguities in the task, missing context, and format drift. Its failures are information for the next prompt.

> **Examples Need Boundaries**
>
> Clear markers around EXAMPLE INPUT, EXAMPLE OUTPUT, and YOUR INPUT reduce confusion between the demonstration and the live task.



AI-generated (gpt-image-1.5)

> **Choose a Representative Case**
>
> An example should exercise the rule you care about. An unusually easy or exceptional case can teach the wrong boundary.



AI-generated (gpt-image-1.5)

Examples Communicate More Than Labels

Demonstrations can reveal output format, available labels, and the distribution of inputs, not only a direct reasoning rule.

> **Curate the Set**
>
> More examples use more context and can introduce contradictions. Keep examples relevant, accurate, and distinct enough to add information.

> **Change One Major Variable**
>
> Holding the test cases steady while changing one instruction or example makes the effect easier to interpret.

> **Diagnose Before Rewriting**
>
> Different failures need different fixes. Missing evidence is not repaired by a more enthusiastic tone, and a wrong format may not need a new model.

Matched pairs

Each item on the left goes with the one on the right.

Zero-shot prompt → Instructions without a worked example

Probe run → A first attempt used to discover failure modes

One-shot prompt → Instructions plus one demonstration

Representative example → A case that exposes the intended rule

Few-shot prompt → Instructions plus several demonstrations

Boundary example → A case near the edge of a category

Prompt regression → A previously good case fails after a change

Test set → Representative inputs used across prompt versions

Step by step

Build a few-shot example set

1. Name the distinction the examples must teach
2. Add one ordinary positive case
3. Add a useful edge case
4. Add a near miss or contrasting case
5. Check labels, format, and instructions for consistency

Prefer a small, coherent set over a long pile of redundant examples.

Run a prompt improvement cycle

1. Save the current prompt and representative inputs
2. Score outputs with a short rubric
3. Name the dominant failure pattern
4. Change one major instruction or example
5. Retest and keep the version only if it improves the set

Record regressions as well as wins; a prompt can improve one case while harming another.

Self-test

Answer each question from memory, then check your work against the answer key.

1. What can a zero-shot prompt include?

- A Only one sentence
- B Context and constraints
- C A trained adapter
- D No source material

2. When is zero-shot a good start?

- A A hidden grading label
- B A safety-critical decision
- C A novel private code
- D A familiar low-risk task

3. What is one-shot prompting?

- A Training one parameter
- B Running the model once
- C Giving one worked example
- D Using one-word answers

4. Why delimit an example?

- A To separate it from live input
- B To guarantee accuracy
- C To reduce all token use
- D To add hidden memory

5. What makes a few-shot set useful?

- (A) Varied boundary coverage
- (B) Hidden instructions
- (C) Maximum repetition
- (D) Random long passages

Go deeper

WATCH The First Prompt Is a Probe 0:31
Hand a model your notes with one instruction: turn these into bullets. No example of the answer.

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/learn-by-showing

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 3 OF 14

Roles, Reasons, and Control

Roles Are Lenses · Ask for Visible Reasons · Break Work into Stages · Surface Assumptions and Uncertainty

WHAT YOU'LL BE ABLE TO DO

- Use roles, decomposition, assumptions, and visible checks without mistaking performance cues for expertise.

Roles Are Lenses

A role prompt can establish audience, vocabulary, priorities, or a review lens: “Act as a skeptical copy editor” is a compact way to request attention to ambiguity and unsupported claims. It is a direction for behavior, not a credential check.

Research on factual question answering found that personas do not reliably improve objective knowledge and their effects vary across models and questions. Put the task, evidence, and rubric first. Use a role only when it changes observable behavior you actually want.

Three ideas to retain:

- **A Role Sets a Lens:** Roles can cue tone, audience, priorities, or review standards. They are most useful when the desired perspective can be described and evaluated.
- **A Persona Is Not a Credential:** “Act as a doctor” does not confer medical training, accountability, or access to a patient's record. The output still needs appropriate evidence and review.
- **Task Details Usually Matter More:** A clear assignment, source packet, and rubric provide stronger control than an impressive job title alone.

Key term: **Role prompt** means “an instruction to adopt a perspective or standard.”

Ask for Visible Reasons

For important work, ask for inspectable support: cite the supplied passage, show the formula, list assumptions, compare alternatives, or provide a concise rationale. Those artifacts help a reviewer locate the step that needs checking.

A long explanation is not automatically faithful to how a model produced its answer, and verbosity can make a weak answer look rigorous. Prefer short, task-relevant evidence over theatrical internal monologues. Verify the underlying claim or calculation independently when the stakes require it.

Three ideas to retain:

- **Request Inspectable Artifacts:** Citations, formulas, assumptions, intermediate tables, and test results give a reviewer concrete objects to check.

- **Length Is Not Proof:** A detailed explanation can still be unsupported, post-hoc, or wrong. Fluency and confidence do not substitute for evidence.
- **Match the Reason to the Task:** A calculation needs shown inputs and arithmetic; a source-based summary needs claim-to-passage links; a recommendation needs criteria and tradeoffs.

Key term: **Concise rationale** means "a short justification tied to evidence."

Break Work into Stages

Complex requests often improve when separated into stages. First inventory the inputs and unknowns, then propose a plan, produce a draft, and finally inspect it against criteria. Staging keeps early misunderstandings from contaminating every downstream step.

Do not decompose forever. Every added step costs time and can create new failure points. Use stages when they expose a decision, produce a checkable artifact, or create a useful approval point. For a trivial task, a direct prompt is usually better.

Three ideas to retain:

- **Stages Expose Decisions:** Planning, drafting, and checking as separate phases makes assumptions and intermediate choices easier to inspect.
- **Approval Can Sit Between Stages:** A human can approve the outline before prose is produced or approve a query before it touches a database.
- **Complexity Has a Cost:** Extra steps add latency and opportunities for drift. Decompose only where a stage creates control or useful evidence.

Key term: **Planning stage** means "chooses an approach before production."

Surface Assumptions and Uncertainty

A model often completes an underspecified request by selecting plausible assumptions. That can be helpful in brainstorming and dangerous in decisions. Ask it to separate supplied facts, inferences, assumptions, and unknowns. Require clarifying questions when a missing answer would materially change the result.

Uncertainty language is not perfectly calibrated, so "I am 80% confident" is not a measurement unless the system has been evaluated for that task. Use uncertainty as a routing signal: which claim needs a source, which choice needs a human, and which missing input blocks progress.

Three ideas to retain:

- **Assumptions Fill Gaps:** When a request leaves key details open, a model may choose plausible values rather than stop. Making assumptions explicit keeps them reviewable.
- **Unknown Is a Valid Output:** A responsible response can identify missing information, state what evidence would resolve it, or decline to make a high-stakes guess.
- **Confidence Needs Calibration:** A numeric confidence claim is useful only when its relationship to actual accuracy has been tested for the task and setting.

Key term: **Assumption** means “a premise used without confirmation.”

Sources

- Zheng et al. — Personas in System Prompts Do Not Improve Performances of Large Language Models (Findings of EMNLP 2024)
- Google Cloud — Prompt design strategies
- NIST — Generative AI Profile

Key terms

Persona prompting
Directing a model to respond from a named role or perspective.

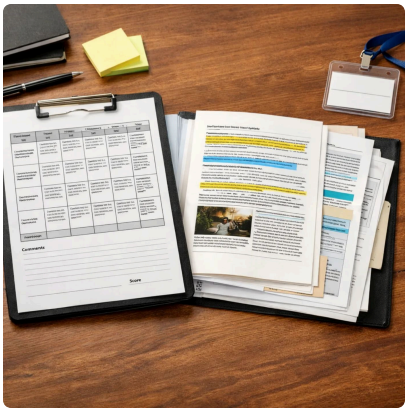
Rationale
A stated justification for a conclusion or choice.

Task decomposition
Dividing a complex job into smaller, inspectable stages.

Calibration
How closely stated confidence corresponds to observed accuracy.

Key facts

> **A Persona Is Not a Credential**
> “Act as a doctor” does not confer medical training, accountability, or access to a patient's record. The output still needs appropriate evidence and review.

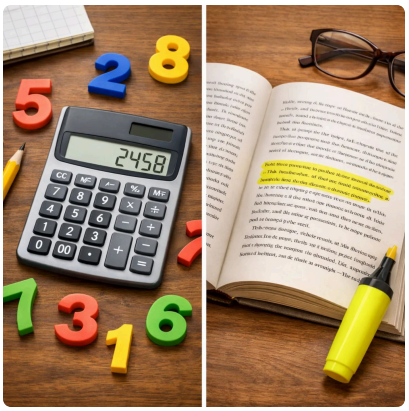


AI-generated (gpt-image-1.5)

Task Details Usually Matter More

A clear assignment, source packet, and rubric provide stronger control than an impressive job title alone.

> **Length Is Not Proof**
> A detailed explanation can still be unsupported, post-hoc, or wrong. Fluency and confidence do not substitute for evidence.



AI-generated (gpt-image-1.5)

Match the Reason to the Task

A calculation needs shown inputs and arithmetic; a source-based summary needs claim-to-passage links; a recommendation needs criteria and tradeoffs.



Wikimedia Commons · Supersonicman13 · CC BY-SA 4.0

Approval Can Sit Between Stages

A human can approve the outline before prose is produced or approve a query before it touches a database.

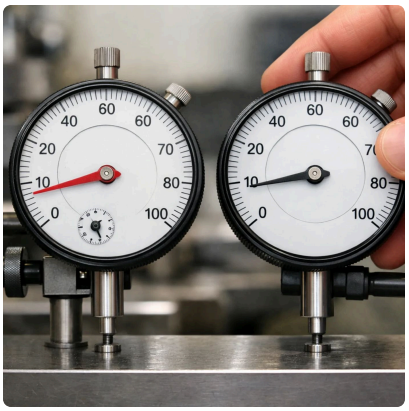
> **Complexity Has a Cost**
>
> Extra steps add latency and opportunities for drift. Decompose only where a stage creates control or useful evidence.



Wikimedia Commons · Kwameghana(Bright Kwame Ayisi) · CC BY-SA 4.0

Unknown Is a Valid Output

A responsible response can identify missing information, state what evidence would resolve it, or decline to make a high-stakes guess.



AI-generated (gpt-image-1.5)

Confidence Needs Calibration

A numeric confidence claim is useful only when its relationship to actual accuracy has been tested for the task and setting.

Matched pairs

Each item on the left goes with the one on the right.

Role prompt → An instruction to adopt a perspective or standard

Expertise claim → A statement of actual competence or qualification

Concise rationale → A short justification tied to evidence

Verifiable trace → An artifact a reviewer can independently check

Planning stage → Chooses an approach before production

Review stage → Checks output against requirements and evidence

Assumption → A premise used without confirmation

Clarifying question → A request for missing information that changes the task

Step by step

Decompose a complex AI-assisted task

1. List the available inputs and unknowns
2. Ask for a plan with decision points
3. Approve or revise the plan
4. Generate one stage at a time
5. Check each deliverable before continuing

Keep a stage only when it reveals a choice, artifact, or approval boundary.

Create an assumption check

1. List facts explicitly supplied by the user
2. List inferences drawn from those facts
3. Mark every unconfirmed premise as an assumption
4. Ask about assumptions that would change the result
5. Route unresolved uncertainty to verification or human review

Do not turn a model's confidence wording into a probability without calibration evidence.

Self-test

Answer each question from memory, then check your work against the answer key.

1. What can a role prompt reliably provide?

- A Guaranteed factual accuracy
- B Private case knowledge
- C A requested perspective
- D A professional license

2. Which role instruction is most testable?

- A Know every hidden fact
- B Review this as a copy editor
- C Never make an error
- D Be the world's best genius

3. Which request is easiest to verify?

- A Show inputs and formula
- B Be completely certain
- C Write a longer answer
- D Think more deeply

4. What does a long explanation prove?

- A That sources were checked
- B That risk is eliminated
- C That every step is true
- D Only that text was produced

5. Why approve an outline first?

- A Guarantee every fact
- B Catch direction errors early
- C Retrain the model
- D Remove all later review

Go deeper

WATCH The Doctor Prompt 0:31
Ask an artificial intelligence model to act as a doctor. Does it become one?

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/roles-reasons-and-control

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 4 OF 14

From Chatbots to Agents

The Agent Loop · Tools, State, and Permissions · Human Approval Checkpoints · Agent Failure Modes

WHAT YOU'LL BE ABLE TO DO

- Design bounded agent workflows with tools, permissions, checkpoints, and explicit stopping conditions.

The Agent Loop

An agent is more than a chat response. It pursues a goal across multiple steps, selects actions or tools, observes results, and decides whether to continue. The surrounding software—not the language model alone—usually manages state, permissions, tool calls, and termination.

A practical loop is plan, act, observe, check, then stop or revise. Every loop needs a definition of done and a limit on attempts, time, or cost. Without those boundaries, a system can repeat a failed action, expand its scope, or consume resources without improving the result.

Three ideas to retain:

- **Agents Operate Across Steps:** A goal-directed agent can select tools, inspect results, update state, and choose a next action instead of waiting for a new user message each time.
- **Software Holds the Boundaries:** The agent harness commonly supplies tools, permissions, memory, logs, and stop logic around the model.
- **Every Loop Needs an Exit:** A success condition plus limits on attempts, time, and cost keeps a stalled workflow from running indefinitely.

Key term: **Agent observation** means “the result returned after an action.”

Tools, State, and Permissions

Tools turn text into action: search, files, calendars, databases, code execution, or messages. State lets the workflow remember what happened. Permissions determine which resources and operations are available. These are product and security decisions, not merely prompt details.

Use least privilege. A research agent may need read access to a folder but not permission to delete it; a draft assistant may prepare an email without sending it. Separate read, create, update, send, and delete capabilities wherever the platform allows.

Three ideas to retain:

- **Tools Create Real Effects:** A model response becomes consequential when connected to search, files, databases, code, purchases, or communication channels.

- **State Records the Journey:** Workflow state can track inputs, actions, observations, approvals, and remaining work across steps.
- **Least Privilege Limits Damage:** Grant only the data and actions needed for the current task, for only as long as needed.

Key term: **Read permission** means "ability to inspect without changing data."

Human Approval Checkpoints

Human-in-the-loop is useful only when the human sees the right information at the right moment. Place approval before money moves, data is deleted, messages are sent, records change, or a high-stakes recommendation reaches a decision maker.

A rubber-stamp screen is not a safeguard. Show the proposed action, relevant evidence, uncertainty, and expected effect; give the reviewer time and authority to reject it. Lower-risk reversible steps can run automatically while consequential steps pause.

Three ideas to retain:

- **Review Belongs Before Consequence:** An approval is most valuable before an action becomes costly, external, legally significant, or difficult to reverse.
- **Reviewers Need Evidence:** A useful checkpoint shows what will happen, why, which inputs support it, and what uncertainty remains.
- **Automation Can Be Tiered:** Low-risk reversible work may proceed automatically while sensitive, destructive, financial, or public actions require explicit approval.

Key term: **Approval gate** means "a required human decision before an action."

Agent Failure Modes

Agents can pursue the wrong interpretation, call the wrong tool, trust hostile content, repeat a failed action, or report success without verifying the result. Because steps compound, a small early mistake can shape every later action.

Design for failure: validate tool inputs and outputs, isolate untrusted content from instructions, log actions, make writes reversible, cap retries, and create an escalation path. Test ordinary failures and adversarial cases before expanding authority.

Three ideas to retain:

- **Errors Compound Across Steps:** A mistaken plan or contaminated observation can influence every later decision in a multi-step workflow.
- **External Content Can Be Hostile:** A webpage, email, or document may contain instructions intended to redirect an agent. Treat retrieved content as data, not trusted control text.
- **Logs Enable Reconstruction:** Action and approval logs help reviewers see what happened, diagnose failure, and improve the workflow.

Key term: **Agent drift** means “movement away from the user's intended goal.”

Sources

- OpenAI — A practical guide to building agents
- NIST — Strengthening AI agent hijacking evaluations
- NIST — Generative AI Profile

Key terms

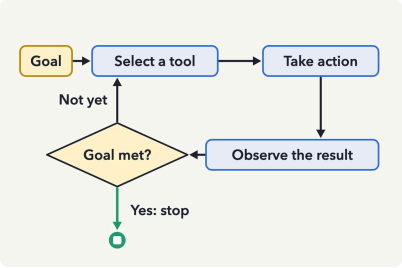
AI agent
A goal-directed system that can choose and execute actions across multiple steps.

Least privilege
Giving a user or system only the minimum access needed for its task.

Human-in-the-loop
A workflow in which a person reviews, directs, or approves selected system actions.

Prompt injection
Untrusted content designed to override or redirect an AI system's instructions.

Key facts



FlashFeed original diagram · CC-BY-4.0

Agents Operate Across Steps

A goal-directed agent can select tools, inspect results, update state, and choose a next action instead of waiting for a new user message each time.

> **Software Holds the Boundaries**
> The agent harness commonly supplies tools, permissions, memory, logs, and stop logic around the model.

> **Every Loop Needs an Exit**
> A success condition plus limits on attempts, time, and cost keeps a stalled workflow from running indefinitely.

> **State Records the Journey**
> Workflow state can track inputs, actions, observations, approvals, and remaining work across steps.

> **Least Privilege Limits Damage**
> Grant only the data and actions needed for the current task, for only as long as needed.



AI-generated (gpt-image-1.5)

Reviewers Need Evidence

A useful checkpoint shows what will happen, why, which inputs support it, and what uncertainty remains.

> **Automation Can Be Tiered**
>
> Low-risk reversible work may proceed automatically while sensitive, destructive, financial, or public actions require explicit approval.

> **External Content Can Be Hostile**
>
> A webpage, email, or document may contain instructions intended to redirect an agent. Treat retrieved content as data, not trusted control text.

> **Logs Enable Reconstruction**
>
> Action and approval logs help reviewers see what happened, diagnose failure, and improve the workflow.

Matched pairs

Each item on the left goes with the one on the right.

Agent observation → The result returned after an action

Stop condition → A rule that ends the agent loop

Read permission → Ability to inspect without changing data

Write permission → Ability to create or modify data

Approval gate → A required human decision before an action

Reversible action → A change that can be safely undone

Agent drift → Movement away from the user's intended goal

Escalation path → A route to human help when the workflow cannot proceed safely

Step by step

Bound an agent loop

1. State the goal and definition of done
2. List allowed tools and actions
3. Set attempt, time, and cost limits
4. Check each observation against the goal
5. Stop, escalate, or revise when a limit is reached

A loop that cannot explain when it should stop is not ready to run unattended.

Set permissions for an agent

1. List every required resource
2. Separate read, create, update, send, and delete actions
3. Remove capabilities the goal does not need
4. Require approval for consequential operations
5. Revoke temporary access when the task ends

Permission boundaries should be enforced by the surrounding system, not only requested in prose.

Test an agent before wider use

1. Run representative normal tasks
2. Simulate missing and malformed tool results
3. Insert hostile instructions in retrieved content
4. Verify logs, limits, and approval gates
5. Expand authority only after failures are contained

Test the control system around the model, not just the quality of its prose.

Self-test

Answer each question from memory, then check your work against the answer key.

1. What most distinguishes an agent?

- A A friendly chat tone
- B Multi-step action toward a goal
- C A long context window
- D A humanlike avatar

2. What prevents an endless loop?

- A An explicit stop condition
- B A hidden greeting
- C A longer role prompt
- D More adjectives

3. What access should a summarizer get?

- A Payment authority
- B Delete access to every folder
- C All company passwords
- D Read access to the chosen files

4. Why separate draft from send?

- A Drafts cannot be reviewed
- B Drafting uses no language
- C Sending has an external effect
- D Sending improves accuracy

5. Which action most needs approval?

- A Sending a payment
- B Renaming a scratch heading
- C Suggesting file tags
- D Sorting local draft notes

Go deeper

WATCH Data, Not Orders 0:27
A webpage may contain instructions meant to redirect an artificial intelligence agent. An email or a document may too.

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/from-chatbots-to-agents

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 5 OF 14

Make with Multiple Media

Prompt an Image · Choose the Canvas · Plan Video as Shots · Work with Audio

WHAT YOU'LL BE ABLE TO DO

- Prompt and review image, video, and audio systems with attention to composition, continuity, consent, and fit.

Prompt an Image

An image prompt works like a compact creative brief. Name the subject and action, then setting, composition, camera viewpoint, light, color, medium, mood, and anything that must be absent. Put the most important visual relationship in plain language rather than a pile of aesthetic keywords.

Generate, inspect, and revise. Change one dimension—framing, palette, subject placement, or texture—while preserving the rest. For identity or product consistency, use reference images where the tool supports them and check permissions for every reference.

Three ideas to retain:

- **Start with the Visual Subject:** State what must appear, what it is doing, and the relationship among important elements before adding style language.
- **Composition Directs Attention:** Viewpoint, framing, scale, subject placement, depth, and negative space determine how the eye moves through an image.
- **Iteration Beats Keyword Piles:** Inspect a result and revise a specific visual dimension. Many fashionable adjectives can compete without clarifying the composition.

Key term: **Image subject** means “the main person, object, or scene.”

Choose the Canvas

Aspect ratio is width divided by height. A square canvas supports balanced product or profile compositions; portrait ratios serve posters and phone screens; landscape ratios fit banners, presentations, and wide scenes. The destination should determine the canvas before you refine the composition.

Changing ratio is not neutral cropping. A subject centered in a square may need to move in a wide banner to leave space for a headline. Generate for the target ratio when possible, protect essential details from edges, and preview the asset inside the real layout.

Three ideas to retain:

- **Aspect Ratio Shapes Composition:** The width-to-height relationship changes where subjects, motion, text space, and visual balance can fit.

- **Design for the Destination:** Choose the canvas for the feed, slide, poster, thumbnail, or banner where the asset will actually appear.
- **Cropping Can Remove Meaning:** Converting one ratio to another may cut faces, hands, products, or context. Preview each required crop rather than assuming it is safe.

Practice by sketching the target placement before generation. Choose the destination, select its canvas, place the subject and negative space, then preview the result in the actual layout instead of trusting an isolated image.

Plan Video as Shots

Video adds time. A useful prompt describes not only what a frame contains but what changes: camera movement, subject action, duration, pacing, and the transition into the next shot. A storyboard or shot list makes those decisions visible before generation.

Continuity remains hard. Lock character details, wardrobe, props, lighting, and screen direction; reuse reference frames when available. Generate short shots, select the best takes, and edit them together rather than asking one long generation to solve every beat.

Three ideas to retain:

- **Video Prompts Describe Change:** Action, camera movement, duration, pacing, and transitions matter because the output unfolds over time.
- **Storyboards Externalize the Plan:** A sequence of frames or shot descriptions lets you inspect narrative beats and composition before generating footage.
- **Continuity Needs Deliberate Control:** Characters, props, lighting, geography, and direction can drift between clips unless they are specified and checked.

Key term: **Shot list** means “an ordered plan of camera setups.”

Work with Audio

Audio systems can transcribe speech, separate speakers, translate, synthesize a voice, generate music, or create sound effects. Each task has different quality checks. A transcript needs names, numbers, and speaker turns reviewed; a synthetic voice needs explicit authority and disclosure appropriate to its use.

Voice can function as identity. Short samples may support convincing impersonation, so never assume familiar sound proves a caller is genuine. Obtain consent for cloning, restrict access to voice assets, and verify urgent requests through a trusted second channel.

Three ideas to retain:

- **Audio Tasks Are Not Interchangeable:** Transcription, speaker labeling, translation, voice synthesis, music, and effects use different inputs and require different reviews.
- **Transcripts Need Human Checks:** Names, numbers, technical terms, accents, noise, and overlapping speakers are common places for speech recognition to fail.

- **A Voice Can Be an Identity Signal:** Synthetic speech can imitate a person, enabling accessibility and creativity as well as impersonation and fraud.

Practice by sorting the six named moves in this section, then choose one real task and explain which move fits it. For consequential work, keep the evidence, permissions, and final judgment with an accountable person.

Sources

- OpenAI Academy — Image generation
- Google Cloud — Configure image aspect ratio
- Rombach et al. — High-Resolution Image Synthesis with Latent Diffusion Models
- Radford et al. — Robust Speech Recognition via Large-Scale Weak Supervision
- FTC — Approaches to address AI-enabled voice cloning

Key terms

Creative brief
A concise statement of a visual project's goal, subject, audience, and constraints.

Storyboard
A sequence of images or panels that plans how a video will unfold.

Key concepts

Transcription

- Turns recorded speech into written text.
- Names and numbers need special review.
- Its output is a transcript.

Diarization

- Works with recordings containing multiple speakers.
- Answers the question of who spoke when.
- It labels speaker turns.

Speech translation

- Changes spoken language for another audience.
- It combines recognition and translation.
- Its goal is speech in another language.

Voice synthesis

- Produces spoken audio from text or controls.
- It may use a generic or imitated voice.
- Consent matters when identity is copied.

Music generation

- Creates rhythmic and tonal audio.
- Prompts may name instrumentation, tempo, and mood.
- Its output is music.

Sound-effect generation

- Creates short nonmusical audio events.
- Prompts name source, environment, and duration.
- Footsteps or a door close are examples.

Key facts

> **Composition Directs Attention**
> Viewpoint, framing, scale, subject placement, depth, and negative space determine how the eye moves through an image.

> **Iteration Beats Keyword Piles**
> Inspect a result and revise a specific visual dimension. Many fashionable adjectives can compete without clarifying the composition.

> **Design for the Destination**
> Choose the canvas for the feed, slide, poster, thumbnail, or banner where the asset will actually appear.

> **Cropping Can Remove Meaning**
> Converting one ratio to another may cut faces, hands, products, or context. Preview each required crop rather than assuming it is safe.

> **Storyboards Externalize the Plan**
> A sequence of frames or shot descriptions lets you inspect narrative beats and composition before generating footage.

> **Continuity Needs Deliberate Control**
> Characters, props, lighting, geography, and direction can drift between clips unless they are specified and checked.

> **Transcripts Need Human Checks**
> Names, numbers, technical terms, accents, noise, and overlapping speakers are common places for speech recognition to fail.

> **A Voice Can Be an Identity Signal**
> Synthetic speech can imitate a person, enabling accessibility and creativity as well as impersonation and fraud.

By the numbers

Square canvas
1 width ÷ height

Portrait 3:4 canvas
0.75 width ÷ height

Landscape 4:3 canvas
1.333 width ÷ height

Vertical 9:16 canvas
0.5625 width ÷ height

Widescreen 16:9 canvas
1.778 width ÷ height

Classic photo 3:2 canvas
1.5 width ÷ height

Matched pairs

Each item on the left goes with the one on the right.

Image subject → The main person, object, or scene

Negative prompt → A description of elements to avoid

Shot list → An ordered plan of camera setups

Continuity → Consistency of visual details across shots

Step by step

Revise a generated image

1. Identify the one visual problem that matters most
2. Name the desired change in composition, light, color, or subject
3. Preserve elements that already work
4. Generate a small set of variants
5. Check details, rights, and intended use before publishing

Keep the revision specific enough that you can tell whether it succeeded.

Plan a generated video sequence

1. Write the audience and single message
2. Break the message into visual beats
3. Specify each shot's subject, action, camera, and duration
4. Lock continuity details and reference frames
5. Generate short takes, review, and edit

Sound, captions, rights, and accessibility belong in the production plan too.

Self-test

Answer each question from memory, then check your work against the answer key.

1. What should an image prompt name first?

A Twenty style buzzwords
B A random camera brand
C An invisible watermark
D The essential subject

2. A portrait feels cramped. Best revision?

A Tighten the camera framing
B Ask for more negative space
C Increase text length
D Add more art styles

3. Which canvas is tallest relative to width?

A 1:1 square
B 4:3 landscape
C 9:16 vertical
D 16:9 widescreen

4. What does a video prompt add beyond a still?

A Change over time
B A single aspect ratio only
C Guaranteed continuity
D Higher image resolution

5. Best way to control a long sequence?

A Ignore screen direction
B Use one vague prompt
C Remove all references
D Generate and edit short shots

Go deeper

WATCH Not a Neutral Crop 0:29
A product sits centered in a square image. In a wide banner, that product may need to move to leave space for a headline.

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/make-with-multiple-media

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 6 OF 14

Writing and Documents

Separate Drafting from Editing · Summarize Long Material · Transform Tone and Format · Ask Questions of Documents

WHAT YOU'LL BE ABLE TO DO

- Use AI to draft, revise, summarize, transform, and inspect documents while preserving meaning and provenance.

Separate Drafting from Editing

Drafting starts from a goal and source material; editing starts from existing language. Name which job you want. “Rewrite this” may invite unnecessary invention, while “flag unclear sentences without changing the text” produces a diagnostic pass you can inspect first.

For substantive revision, state what must be preserved: facts, voice, quotations, dates, commitments, or legal terms. Ask for a change log when silent alterations would matter. Keep the original beside the revision so fluency does not hide lost meaning.

Three ideas to retain:

- **Drafting and Editing Are Different Jobs:** Drafting creates a first version from a brief. Editing diagnoses or changes existing language according to stated goals.
- **Preservation Rules Prevent Drift:** Identify facts, quotations, commitments, voice, and other elements that must survive a rewrite.
- **A Change Log Makes Revision Visible:** For consequential documents, a short list of material edits helps the author inspect what changed and why.

Key term: **Developmental edit** means “changes structure, argument, or coverage.”

Summarize Long Material

A summary is a selection, not a smaller copy. Specify who needs it, which questions it should answer, what must be retained, and the desired length. Distinguish an executive summary, abstract, meeting recap, chronology, and extractive quotation list.

Long inputs create special risks: decisive details can be missed, especially when buried among many documents. Chunk with overlap when needed, summarize sections first, and build a final synthesis from those traceable notes. Require page, section, timestamp, or passage references for material claims.

Three ideas to retain:

- **A Summary Has a Purpose:** Audience and decision need determine which details are central, which are supporting, and which can be omitted.

- **Long Context Is Not Perfect Recall:** Research finds that model performance can vary with where relevant information appears, including degradation for material in the middle.
- **Traceability Improves Review:** Section, page, timestamp, or passage links let a reader compare summary claims with the original material.

Key term: **Executive summary** means "decision-focused overview for a busy reader."

Transform Tone and Format

AI is useful for turning notes into an email, prose into a table, or a technical explanation into plain language. A transformation should state the new audience and format plus the semantic invariants—the facts, qualifications, quantities, and commitments that cannot change.

Tone is relational, not decorative. "Warm" for a condolence note differs from "warm" for a sales message. Describe the relationship and purpose, then inspect whether the revision added promises, removed caveats, or sharpened uncertain language into certainty.

Three ideas to retain:

- **Transformation Changes Representation:** A format or tone change should preserve the underlying facts and intended meaning unless the brief explicitly requests substantive revision.
- **Audience Defines Plain Language:** Useful simplification depends on what the reader already knows and what action they need to take.
- **Fluency Can Hide Semantic Drift:** A smoother rewrite may drop a caveat, change a quantity, or turn a possibility into a promise.

Key term: **Semantic invariant** means "meaning that must remain unchanged."

Ask Questions of Documents

Document question answering works best when the answer must come from a defined packet. Tell the system to cite the exact passage, say when the packet does not answer the question, and keep external knowledge separate. For extraction, define a schema and what to do with missing fields.

A missing mention is not proof that an event never occurred. "Not found in these documents" is narrower and more honest than "false." Sample the extraction against originals, especially dates, names, negation, units, and table cells.

Three ideas to retain:

- **Bound the Evidence Set:** A document Q&A task should identify which files count as evidence and whether outside sources are allowed.
- **Not Found Is Not False:** Failure to locate a claim in a packet shows only that it was not found there; it does not establish that the claim is untrue everywhere.
- **Schemas Make Extraction Testable:** Named fields, types, and missing-value rules turn loose reading into output that can be sampled and compared with originals.

Key term: **Evidence packet** means “the defined documents allowed to support an answer.”

Sources

- Google Cloud — Prompt design strategies
- Liu et al. — Lost in the Middle
- NIST — Generative AI Profile

Key terms

Revision brief
Instructions describing what to change, what to preserve, and how the result will be judged.

Abstractive summary
A condensed restatement written in new language rather than copied passages.

Tone
The attitude and relationship conveyed by choices of wording, rhythm, and emphasis.

Document grounding
Constraining an answer to information found in a supplied set of documents.

Key facts

> **Preservation Rules Prevent Drift**
> Identify facts, quotations, commitments, voice, and other elements that must survive a rewrite.

> **A Change Log Makes Revision Visible**
> For consequential documents, a short list of material edits helps the author inspect what changed and why.



AI-generated (gpt-image-1.5)

> **Long Context Is Not Perfect Recall**
> Research finds that model performance can vary with where relevant information appears, including degradation for material in the middle.

> **Traceability Improves Review**
> Section, page, timestamp, or passage links let a reader compare summary claims with the original material.



AI-generated (gpt-image-1.5)

Audience Defines Plain Language

Useful simplification depends on what the reader already knows and what action they need to take.



AI-generated (gpt-image-1.5)

Fluency Can Hide Semantic Drift

A smoother rewrite may drop a caveat, change a quantity, or turn a possibility into a promise.

> **Not Found Is Not False**
>
> Failure to locate a claim in a packet shows only that it was not found there; it does not establish that the claim is untrue everywhere.



AI-generated (gpt-image-1.5)

Schemas Make Extraction Testable

Named fields, types, and missing-value rules turn loose reading into output that can be sampled and compared with originals.

Matched pairs

Each item on the left goes with the one on the right.

Developmental edit → Changes structure, argument, or coverage

Copy edit → Improves clarity, consistency, and correctness

Change log → A concise record of material revisions

Executive summary → Decision-focused overview for a busy reader

Extractive summary → A selection of passages taken from the source

Semantic invariant → Meaning that must remain unchanged

Format transformation → Recasting content into a new structure

Plain-language rewrite → A version adapted for readers without specialist knowledge

Evidence packet → The defined documents allowed to support an answer

Missing value rule → How absent or unknown fields should be represented

Source citation → A pointer from an answer to supporting evidence

Self-test

Answer each question from memory, then check your work against the answer key.

1. What should a rewrite prompt protect?

- A Meaning that must not change
- B Only the word count
- C A random persona
- D Every original typo

2. Best first step for sensitive prose?

- A Remove the original
- B Replace it without review
- C Hide all changes
- D Ask for diagnosis before rewriting

3. What should a summary prompt specify?

- A A fictional credential
- B Audience and coverage
- C No source boundaries
- D Only “make shorter”

4. How can a long-document summary be checked?

- A Remove section labels
- B Hide omitted material
- C Link claims to source locations
- D Trust confident wording

5. What must a format conversion preserve?

- A Material facts and qualifications
- B A random metaphor
- C The model's first guess
- D Every line break

Go deeper

WATCH The Buried Detail 0:31
Bury a decisive detail among many documents. Ask an artificial intelligence model for a summary. That detail can be missed.

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/writing-and-documents

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 7 OF 14

Research, Learning, and Decisions

Plan Research Before Answering · Use AI as a Practice Partner · Compare Options without Outsourcing Choice · Six Thinking Moves

WHAT YOU'LL BE ABLE TO DO

- Use AI to structure inquiry, practice learning, compare options, and perform distinct thinking moves without outsourcing judgment.

Plan Research Before Answering

AI can help convert a broad question into subquestions, search terms, inclusion criteria, and an evidence table. Start by asking what would count as an answer and which claims require primary, current, or local sources. This creates a plan you can inspect before reading polished prose.

Generated citations may be invented, misquoted, or merely adjacent to the claim. Open every material source, confirm authorship and date, and check that the cited passage actually supports the sentence. Treat summaries of sources as leads until verified.

Three ideas to retain:

- **Research Starts with Claim Types:** A historical date, medical recommendation, product price, and personal preference require different sources and levels of verification.
- **Search Terms Are Hypotheses:** Suggested queries help explore a topic, but wording can bias what appears. Try synonyms, counterarguments, and source-specific searches.
- **A Citation Must Support the Claim:** A real link is not sufficient if its content does not entail, qualify, or directly document the sentence it is attached to.

Key term: **Primary source** means "original evidence or an authoritative first-party record."

Use AI as a Practice Partner

A tutor is most useful when it makes the learner do cognitive work. Ask for one question at a time, wait for your attempt, request a hint before a solution, and have the system diagnose the misconception. Use retrieval practice and varied examples rather than rereading a generated explanation.

The learner still needs trusted source material and human support where appropriate. AI feedback may be persuasive and wrong. Compare explanations with the course text, teacher guidance, worked solutions, or another authoritative source, and do not upload protected student data casually.

Three ideas to retain:

- **Practice Beats Passive Answering:** Questions, retrieval, explanation, and correction require the learner to produce knowledge instead of only reading a generated solution.

- **Hints Preserve Productive Struggle:** A staged hint can unblock progress without immediately replacing the learner's reasoning.
- **Tutoring Output Needs a Reference:** A model can give confident but incorrect feedback, so learning sessions should connect to a trusted text, teacher, or answer key.

Key term: **Retrieval practice** means “recalling knowledge without looking at the answer.”

Compare Options without Outsourcing Choice

AI can organize options, suggest criteria, surface tradeoffs, and run scenarios. It should not quietly choose your values. State which outcomes matter, distinguish must-haves from preferences, and provide evidence for scores instead of accepting invented numbers.

Use sensitivity analysis: change a weight or assumption and see whether the ranking changes. A fragile recommendation deserves more investigation. Keep irreversible, ethical, legal, medical, and financial decisions with appropriately qualified humans.

Three ideas to retain:

- **Criteria Encode Values:** A decision table reflects what the chooser counts, how it is weighted, and which tradeoffs are acceptable.
- **Scores Need Evidence:** A neat matrix can manufacture precision if ratings are guesses. Link each material score to data, experience, or an explicit assumption.
- **Sensitivity Reveals Fragility:** If a small change in one weight reverses the recommendation, the decision depends heavily on that judgment.

Key term: **Must-have criterion** means “a requirement an option cannot trade away.”

Six Thinking Moves

“Help me think” hides several distinct jobs. You may need expansion, compression, comparison, critique, reframing, or sequencing. Naming the move gives the model a clearer role and helps you judge whether it was useful.

Move deliberately between modes. Brainstorm without judging, then cluster and compare; critique a chosen option, then plan a test. Do not let a long list impersonate a decision. The human supplies purpose, taste, responsibility, and the final commitment.

Three ideas to retain:

- **Thinking Has Modes:** Generating possibilities, reducing information, comparing alternatives, challenging claims, reframing problems, and sequencing action are different tasks.
- **Separate Expansion from Selection:** Brainstorming and judging at the same time can narrow the search too early. Generate first, then apply criteria.
- **Lists Are Not Decisions:** AI can widen and organize a choice set, but a person remains responsible for values, evidence, and final commitment.

Practice by sorting the six named moves in this section, then choose one real task and explain which move fits it. For consequential work, keep the evidence, permissions, and final judgment with an accountable person.

Sources

- NIST — Generative AI Profile
- UNESCO — Guidance for generative AI in education and research
- Liu et al. — Lost in the Middle

Key terms

Evidence table

A structured record linking claims to sources, passages, dates, and confidence notes.

Formative feedback

Information intended to improve learning during practice rather than only score the final result.

Decision matrix

A table that compares options across explicitly chosen criteria.

Key concepts

Brainstorming

- Expands the option space before judgment.
- Quantity and variety matter at first.
- It generates possibilities.

Summarizing

- Compresses a larger body of material.
- It selects what matters for an audience.
- It creates a shorter representation.

Comparing

- Places alternatives against shared criteria.
- It makes similarities and differences visible.
- It examines options side by side.

Critiquing

- Searches for weaknesses, omissions, and counterarguments.
- It pressure-tests an idea.
- Its job is constructive challenge.

Reframing

- Changes the question or point of view.
- It can reveal a hidden assumption.
- It offers a different way to define the problem.

Sequencing

- Turns a goal into an ordered path.
- Dependencies and checkpoints matter.
- It arranges actions over time.

Key facts

> **Search Terms Are Hypotheses**
> Suggested queries help explore a topic, but wording can bias what appears. Try synonyms, counterarguments, and source-specific searches.

> **A Citation Must Support the Claim**
> A real link is not sufficient if its content does not entail, qualify, or directly document the sentence it is attached to.

> **Hints Preserve Productive Struggle**
> A staged hint can unblock progress without immediately replacing the learner's reasoning.

> **Tutoring Output Needs a Reference**
> A model can give confident but incorrect feedback, so learning sessions should connect to a trusted text, teacher, or answer key.

> **Scores Need Evidence**
> A neat matrix can manufacture precision if ratings are guesses. Link each material score to data, experience, or an explicit assumption.

> **Sensitivity Reveals Fragility**
> If a small change in one weight reverses the recommendation, the decision depends heavily on that judgment.

> **Separate Expansion from Selection**
> Brainstorming and judging at the same time can narrow the search too early. Generate first, then apply criteria.

> **Lists Are Not Decisions**
> AI can widen and organize a choice set, but a person remains responsible for values, evidence, and final commitment.

Matched pairs

Each item on the left goes with the one on the right.

Primary source → Original evidence or an authoritative first-party record

Secondary source → Analysis or synthesis based on other sources

Claim-source alignment → The source directly supports the attached statement

Retrieval practice → Recalling knowledge without looking at the answer

Scaffolded hint → A hint that preserves the learner's next step

Socratic question → A prompt that advances thinking without supplying the conclusion

Must-have criterion → A requirement an option cannot trade away

Sensitivity analysis → Testing how conclusions change when assumptions change

Decision owner → The person accountable for the final choice

Self-test

Answer each question from memory, then check your work against the answer key.

1. What should happen to an AI citation?

A Count the URL as proof
B Assume the quote exists
C Open and verify it
D Trust the title alone

2. Why search counterarguments?

A Guarantee consensus
B Reduce one-sided evidence
C Avoid primary sources
D Make every answer longer

3. Which tutor prompt keeps the learner active?

A Wait for my attempt first
B Skip feedback
C Remove all questions
D Give every answer immediately

4. How should AI feedback be treated?

A Assume confidence means truth
B Replace the course source
C Hide disagreements
D Check it against trusted material

5. What makes a decision matrix deceptive?

A Visible tradeoffs
B Source links
C Unsupported precise scores
D Named criteria

6. Why name the specific thinking move (e.g., critique, reframe) you want AI to do?

- A Guarantees the AI never makes a factual error
- B Lets the AI decide your final values for you
- C Removes the need to check citations afterward
- D Gives the model a clearer role you can judge

7. Why keep brainstorming separate from judging ideas early on?

- A Early judging can narrow the search too soon
- B Selection criteria are invented after listing ends
- C Brainstorming alone produces a final decision
- D Expansion and selection are one identical task

Go deeper

WATCH The Page Behind the Link 0:30
An answer from artificial intelligence cites a source. The link is real. The page behind it may still not support the claim.

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/research-learning-and-decisions

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 8 OF 14

Data, Spreadsheets, and Code

Describe the Data First · Build Formulas and Charts · Use a Coding Cycle · Test and Secure Generated Code

WHAT YOU'LL BE ABLE TO DO

- Use AI to inspect structured data and code through schemas, tests, security checks, and reversible changes.

Describe the Data First

A table is not self-explanatory. Column names can be ambiguous, units may differ, categories may encode missing values, and a row may represent a person, transaction, day, or device. Supply a data dictionary and state the question before asking for analysis.

Profile the data first: row count, types, ranges, missingness, duplicates, and impossible values. Preserve the original file and perform transformations on a copy. A plausible chart built on misread units is still wrong.

Three ideas to retain:

- **Rows Need a Unit of Observation:** State what one row represents. Confusing a transaction with a customer or a reading with a day changes every calculation.
- **A Data Dictionary Prevents Guessing:** Field definitions, units, allowed values, and missing markers give both the model and reviewer a shared interpretation.
- **Profile Before Analysis:** Counts, types, ranges, missingness, duplicates, and invalid values reveal whether a dataset is ready for a question.

Key term: **Data dictionary** means “definitions and rules for dataset fields.”

Build Formulas and Charts

For spreadsheet help, provide column names, a few representative rows, the intended result, and the exact spreadsheet dialect if it matters. Ask the model to explain references and edge cases, then test the formula on hand-computed examples.

A chart answers a question through choices: measure, aggregation, comparison group, time interval, axis, and handling of missing values. Ask why the chart type fits the question. Inspect labels and scales; a polished visual can still aggregate the wrong field.

Three ideas to retain:

- **Formulas Depend on Context:** Cell locations, table names, locale, date formats, blank handling, and software dialect can change the correct expression.

- **Tiny Test Cases Catch Big Errors:** A few rows with known outputs can reveal off-by-one ranges, wrong absolute references, and incorrect treatment of blanks.
- **Charts Encode Analytical Choices:** Aggregation, denominator, interval, axis scale, and grouping determine what a visual comparison actually means.

Key term: **Absolute reference** means “a spreadsheet reference fixed when copied.”

Use a Coding Cycle

Useful coding assistance begins with the actual environment: language and versions, repository conventions, relevant files, existing tests, and the desired behavior. Ask for a plan before a broad change and keep the patch narrow enough to review.

Run the software's own formatter, type checker, tests, and build. Read the diff rather than accepting a successful command as proof. Generated code may be functional on a happy path while mishandling errors, concurrency, security, or integration.

Three ideas to retain:

- **Repository Context Matters:** Correct code depends on versions, APIs, conventions, neighboring modules, and tests that are not visible in an isolated snippet.
- **Small Diffs Are Easier to Trust:** A narrowly scoped patch is simpler to review, test, revert, and compare with the requested behavior.
- **Execution Provides Evidence:** Formatting, types, tests, builds, and targeted runtime checks reveal problems that plausible-looking code can hide.

Key term: **Patch** means “a focused set of code changes.”

Test and Secure Generated Code

Generated code can contain ordinary bugs and security vulnerabilities. Ask for tests covering normal, boundary, malformed, and unauthorized inputs, then run them independently. Review data flows, authentication, secrets, dependencies, and error handling.

Do not execute unfamiliar commands with broad permissions or paste credentials into a model. Use a sandbox for uncertain code, pin and audit dependencies, and follow the project's secure development process. A model's own security review is one signal, not an independent control.

Three ideas to retain:

- **Generated Code Is Untrusted Input:** Fluent code should enter the same review, testing, dependency, and security process as code from any other untrusted source.
- **Happy Paths Are Not Enough:** Boundary, malformed, failure, concurrent, and unauthorized cases reveal behavior that a normal example never exercises.
- **Security Needs Independent Controls:** Sandboxing, least privilege, secret management, dependency review, and human inspection cannot be replaced by asking the same model to approve its output.

Key term: **Sandbox** means “an isolated environment that limits a program's effects.”

Sources

- NIST — Secure Software Development Framework
- Lian et al. — A.S.E. security benchmark for AI-generated code
- NIST — Generative AI Profile

Key terms

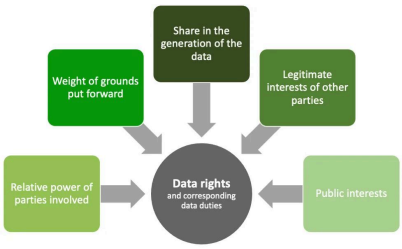
Schema
A structured description of fields, types, relationships, and constraints.

Data validation
Rules that restrict or check values entered into a field or cell.

Code review
Systematic inspection of a change for correctness, clarity, security, and fit.

Threat model
A structured account of assets, attackers, entry points, and plausible harms.

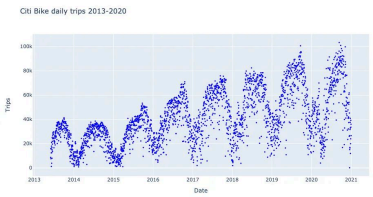
Key facts



Wikimedia Commons · Christiane Wendehorst · CC BY-SA 4.0

A Data Dictionary Prevents Guessing

Field definitions, units, allowed values, and missing markers give both the model and reviewer a shared interpretation.



Wikimedia Commons · Stanislav Rogozhin · Public domain

Profile Before Analysis

Counts, types, ranges, missingness, duplicates, and invalid values reveal whether a dataset is ready for a question.

> **Tiny Test Cases Catch Big Errors**
>
> A few rows with known outputs can reveal off-by-one ranges, wrong absolute references, and incorrect treatment of blanks.

Charts Encode Analytical Choices

Aggregation, denominator, interval, axis scale, and grouping determine what a visual comparison actually means.

Small Diffs Are Easier to Trust

A narrowly scoped patch is simpler to review, test, revert, and compare with the requested behavior.

Execution Provides Evidence

Formatting, types, tests, builds, and targeted runtime checks reveal problems that plausible-looking code can hide.

Happy Paths Are Not Enough

Boundary, malformed, failure, concurrent, and unauthorized cases reveal behavior that a normal example never exercises.



Wikimedia Commons · Mmasontech · CC BY-SA 4.0

Security Needs Independent Controls

Sandboxing, least privilege, secret management, dependency review, and human inspection cannot be replaced by asking the same model to approve its output.

Matched pairs

Each item on the left goes with the one on the right.

Data dictionary → Definitions and rules for dataset fields

Unit of observation → The entity or event represented by one row

Missingness rate → The share of values absent in a field

Absolute reference → A spreadsheet reference fixed when copied

Aggregation → Combining values with a rule such as sum or average

Chart annotation → Text that explains a notable value or event

Patch → A focused set of code changes

Regression test → A test that prevents a fixed bug from returning

Static check → Analysis performed without running the program

Sandbox → An isolated environment that limits a program's effects

Secret → Sensitive credential such as a token or private key

Dependency audit → A check of third-party packages and known risks

Self-test

Answer each question from memory, then check your work against the answer key.

1. What should be stated about each row?

- A The model's persona
- B A hidden conclusion
- C What it represents
- D A preferred chart color

2. What comes before a trend claim?

- A Invent missing values
- B Choose the conclusion
- C Delete the original
- D Profile and clean the data

3. How should an AI formula be checked?

- A Trust syntax alone
- B Test known small cases
- C Hide the inputs
- D Delete the source data

4. What should a chart request identify?

- A Question and aggregation
- B Only a color palette
- C A fictional expert
- D No units

5. What should precede a broad code edit?

- A Rewrite every file
- B Inspect context and plan
- C Assume the API version
- D Remove the tests

Go deeper

WATCH One Signal 0:30
Ask an artificial intelligence model to check its own code for security. Is that enough?

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/data-spreadsheets-and-code

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 9 OF 14

Verify and Protect

Verify Claims and Citations · Minimize Sensitive Data · Resist Injection and Impersonation · Know the High-Stakes Boundary

WHAT YOU'LL BE ABLE TO DO

- Verify material claims, minimize sensitive data, resist manipulation, and keep high-stakes decisions with qualified people.

Verify Claims and Citations

Verification begins by identifying what matters. Separate claims, calculations, quotations, citations, and recommendations; then choose the right check for each. Open primary or authoritative sources, compare exact language, recalculate numbers, and look for omitted qualifications.

Do not ask the same unsupported conversation to certify itself. A second prompt may catch an inconsistency, but independence comes from external evidence, tools, or qualified review. Record what was checked and what remains uncertain.

Three ideas to retain:

- **Verify the Material Claims:** Focus checking effort on statements that drive a decision, create risk, or would be costly to repeat if wrong.
- **Existence and Support Are Separate:** A citation can be real while failing to support the attached claim, or support only a narrower qualified version.
- **Independent Evidence Matters:** Rephrasing the question to the same model can be useful, but authoritative sources, calculations, and external tests provide stronger confirmation.

Key term: **Claim verification** means “checking a statement against appropriate evidence.”

Minimize Sensitive Data

Before entering data, ask whether the task can be completed with less. Remove names, account numbers, health details, private messages, customer records, children's information, and confidential work unless use is authorized and protected by an appropriate service and agreement.

Review the provider's current retention, training, sharing, deletion, security, and administrator controls; these differ by product, plan, region, and settings. Use placeholders or synthetic examples for exploration. Data minimization is useful even when a system claims strong protection.

Three ideas to retain:

- **Data Minimization Reduces Exposure:** Provide only the fields and detail required for the task, and remove direct or indirect identifiers when they are unnecessary.

- **Policies Depend on Product and Setting:** Retention, training use, human review, sharing, deletion, and administrator controls can vary across services and plans.
- **Other People's Data Is Not Yours to Volunteer:** Private messages, records, images, and voices can affect people who did not choose to use the system.

Key term: **Personal data** means “information linked or linkable to a person.”

Resist Injection and Impersonation

An AI system may read webpages, emails, or documents containing instructions aimed at the system rather than the user. Keep control instructions separate from retrieved content, restrict tools, and require confirmation before external actions. Do not let a document grant itself authority.

Synthetic text, images, and voices can make scams more convincing. Urgency, secrecy, payment demands, and channel switching are warning signs regardless of how familiar the voice sounds. Pause and verify through a known number, account, or person.

Three ideas to retain:

- **Retrieved Content Is Untrusted:** Text found in a file or webpage can contain manipulative instructions and should not outrank the workflow's trusted controls.
- **Familiar Sound Is Not Authentication:** Voice cloning can imitate a person well enough to support an emergency or payment scam.
- **Verify through a Trusted Channel:** Use a known phone number, account, contact, or agreed code rather than contact information supplied in the suspicious message.

Key term: **Prompt injection** means “untrusted text that tries to redirect an ai workflow.”

Know the High-Stakes Boundary

Health, legal rights, safety, credit, employment, insurance, and major financial choices can cause serious harm. AI can help explain vocabulary, organize records, draft questions, or compare a document with a checklist. It should not impersonate a qualified professional or make the final decision.

Use authoritative, current, jurisdiction-specific sources and disclose uncertainty. Escalate emergencies immediately to appropriate local services. If an output affects another person's rights or access, require accountable human review and a way to challenge or correct the decision.

Three ideas to retain:

- **High Stakes Change the Standard:** The greater the possible harm, the stronger the evidence, expertise, privacy, review, and accountability that are required.
- **Preparation Is a Safer Role:** AI can help assemble questions, timelines, and source excerpts while a qualified person retains diagnosis, advice, or decision authority.
- **Affected People Need Recourse:** When an AI-assisted decision changes access, opportunity, or rights, people need notice, human review, and a correction or appeal path.

Key term: **Decision support** means “information that helps a responsible person decide.”

Sources

- NIST — Generative AI Profile
- Office of the Privacy Commissioner of Canada — AI chatbots
- FTC — Scammers use AI to enhance family-emergency schemes
- NIST — Strengthening AI agent hijacking evaluations

Key terms

Corroboration
Independent evidence that supports or confirms a claim.

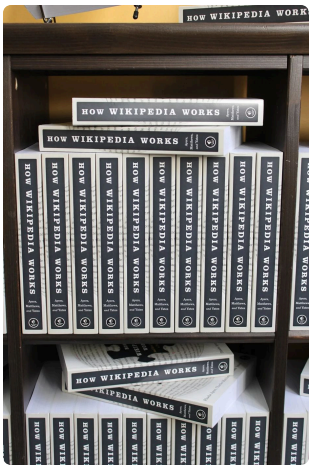
De-identification
Reducing the link between data and the person it describes.

Social engineering
Manipulating a person into revealing information or taking an unsafe action.

High-stakes use
An application where errors can materially affect safety, rights, health, livelihood, or access.

Key facts

> **Existence and Support Are Separate**
>
> A citation can be real while failing to support the attached claim, or support only a narrower qualified version.



Wikimedia Commons · Jay Walsh for the Wikimedia Foundation · CC BY-SA 4.0

Independent Evidence Matters

Rephrasing the question to the same model can be useful, but authoritative sources, calculations, and external tests provide stronger confirmation.

Policies Depend on Product and Setting

Retention, training use, human review, sharing, deletion, and administrator controls can vary across services and plans.

Other People's Data Is Not Yours to Volunteer

Private messages, records, images, and voices can affect people who did not choose to use the system.

Familiar Sound Is Not Authentication

Voice cloning can imitate a person well enough to support an emergency or payment scam.

Verify through a Trusted Channel

Use a known phone number, account, contact, or agreed code rather than contact information supplied in the suspicious message.

Preparation Is a Safer Role

AI can help assemble questions, timelines, and source excerpts while a qualified person retains diagnosis, advice, or decision authority.

Affected People Need Recourse

When an AI-assisted decision changes access, opportunity, or rights, people need notice, human review, and a correction or appeal path.

Matched pairs

Each item on the left goes with the one on the right.

Claim verification → Checking a statement against appropriate evidence

Citation alignment → Whether a source actually supports the attached claim

Personal data → Information linked or linkable to a person

Data minimization → Limiting collection and sharing to what is necessary

Prompt injection → Untrusted text that tries to redirect an AI workflow

Out-of-band check → Verification through a separate trusted channel

Urgency cue → Pressure designed to prevent careful verification

Decision support → Information that helps a responsible person decide

Decision authority → Power and accountability to make the final choice

Qualified reviewer → A person with relevant expertise and accountability

Step by step

Minimize data before using an AI tool

1. Identify the task's essential fields
2. Remove names and unnecessary identifiers
3. Replace real values with placeholders when possible
4. Confirm authorization, policy, and account settings
5. Delete temporary inputs and outputs according to policy

De-identification can fail when remaining details still point to a person.

Self-test

Answer each question from memory, then check your work against the answer key.

1. What must be checked about a citation?

- A That the URL looks long
- B That the title sounds relevant
- C That it supports the claim
- D That the answer is confident

2. Which check is most independent?

- A Repeat the same prompt
- B Ask for more confidence
- C Request a longer rationale
- D Recalculate from source data

3. Best way to test a sensitive workflow?

- A Share every identifier
- B Use synthetic sample data
- C Paste real customer records
- D Disable access controls

4. What should be checked before upload?

- A Current data-handling terms
- B Only the interface color
- C The model's tone
- D A marketing slogan

5. A familiar voice requests urgent money. First move?

- A Share an account password
- B Trust the caller ID
- C Verify through a known channel
- D Pay before the call ends

Go deeper

WATCH A Familiar Voice 0:27
A familiar voice says there is an emergency and asks for money, fast. Is it really them?

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/verify-and-protect

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 10 OF 14

Build Your Personal AI Workflow

Choose the Tool for the Task · Turn Success into a Template · Use a Small Evaluation Set · Practice, Then Go Deeper

> WHAT YOU'LL BE ABLE TO DO
>
> - Choose tools deliberately, turn successful prompts into templates, evaluate them on real cases, and route deeper learning through the collection.

Choose the Tool for the Task

Start with the task, not the brand. Identify the input and output modalities, need for current information or private files, acceptable error, collaboration needs, integrations, latency, and budget. Then compare tools using representative tests.

Selection also includes trust boundaries: retention and training terms, permissions, export and deletion, administrator controls, logs, and human approval. A powerful tool can be the wrong tool when its data handling or action authority exceeds the task.

Three ideas to retain:

- **Task Fit Comes Before Popularity:** The best tool is the one that meets the real input, output, quality, control, and workflow requirements—not necessarily the most famous model.
- **Capabilities Include Controls:** Permissions, logs, deletion, export, administrator settings, and approval gates are part of a tool's practical capability.
- **Run Representative Tests:** A short evaluation set from your real work reveals quality, speed, failure patterns, and cost better than a polished demo.

Key term: **Task fit** means “how well a tool matches a defined job.”

Turn Success into a Template

When a prompt works repeatedly, separate stable instructions from variables. A template can include purpose, input placeholders, constraints, output schema, a demonstration, and a review checklist. Label every field so another person can use it without reverse-engineering your conversation.

Version the template and record what changed. Do not preserve accidental phrases as superstition; keep only components that earn their place in testing. Include an owner, a last-reviewed date, and a retirement condition when the workflow matters.

Three ideas to retain:

- **Templates Separate Stable and Variable Parts:** Purpose, process, and format may stay fixed while audience, source text, dates, and options change each run.

- **Placeholders Need Instructions:** A label such as {SOURCE_TEXT} should say what belongs there, acceptable length, and what to do when the input is missing.
- **Versioning Supports Learning:** A change note and test results preserve why a template evolved and make regressions easier to diagnose.

Key term: **Template variable** means "a field replaced for each run."

Use a Small Evaluation Set

An evaluation set is a handful of representative cases with expected properties. Include ordinary examples, edge cases, and known failure modes. Score dimensions that matter—correctness, completeness, tone, citation support, privacy, latency, and cost—rather than collapsing everything into one vibe.

Set minimum thresholds and identify failures that should stop deployment even if the average is high. Re-run the set after changing a prompt, model, tool, or data source. Production feedback can add new cases, but protect private information when saving them.

Three ideas to retain:

- **Representative Cases Beat One Demo:** A small set spanning normal, edge, and failure conditions provides a more stable signal than one memorable success.
- **Severity Matters beside Frequency:** A rare harmful error may deserve more attention than many harmless formatting misses.
- **Changes Require Retesting:** New prompts, models, tools, integrations, or data can alter performance and revive old failures.

Key term: **Evaluation case** means "an input with expected properties used for testing."

Practice, Then Go Deeper

Choose one recurring, reversible task and run a two-week practice loop. Record the original time, the AI-assisted time, corrections required, sensitive data involved, and whether the result was genuinely useful. Keep the workflow only when evidence justifies the habit.

This pack owns practical technique. Continue to **LLM 101** for tokens, transformers, training, retrieval, and model mechanics. Use **AI for Small Business** for team rollout, customer workflows, governance, and measured pilots. Return to **AI Start Here** for collection vocabulary and debates, and use **The Infinity Machine** for institutional history.

Three ideas to retain:

- **Start with Reversible Work:** A bounded task such as outlining, reformatting, or practice questions lets you learn without giving the system costly authority.
- **Measure Corrections, Not Just Speed:** Time saved before review can disappear when errors, rewriting, privacy cleanup, or rework are counted.

- **Each Learn AI Pack Has a Boundary:** Practical technique belongs here; mechanics, business rollout, orientation, and institutional history have dedicated sibling packs.

Key term: **Bounded pilot** means “a limited trial with success and stop criteria.”

Sources

- NIST — AI Risk Management Framework Playbook
- NIST — Generative AI Profile
- Google Cloud — Prompt design strategies

Key terms

Tool evaluation
A structured comparison of products against task, quality, risk, control, and cost criteria.

Prompt template
A reusable prompt structure with named variables and consistent instructions.

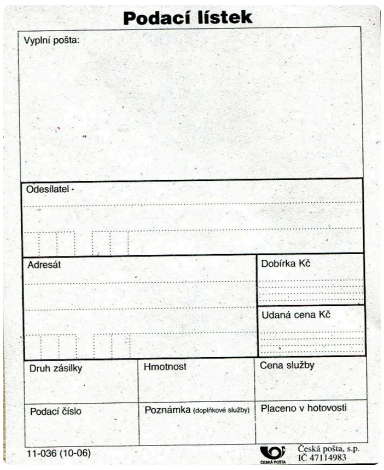
Evaluation rubric
A set of dimensions and standards used to score outputs consistently.

Reversible workflow
A process whose outputs or changes can be safely reviewed and undone.

Key facts

> **Capabilities Include Controls**
>
> Permissions, logs, deletion, export, administrator settings, and approval gates are part of a tool's practical capability.

> **Run Representative Tests**
>
> A short evaluation set from your real work reveals quality, speed, failure patterns, and cost better than a polished demo.



Wikimedia Commons · Pejsek10 · CC BY-SA 3.0

Placeholders Need Instructions

A label such as {SOURCE_TEXT} should say what belongs there, acceptable length, and what to do when the input is missing.

> **Versioning Supports Learning**
>
> A change note and test results preserve why a template evolved and make regressions easier to diagnose.

> **Severity Matters beside Frequency**
>
> A rare harmful error may deserve more attention than many harmless formatting misses.

> **Changes Require Retesting**
>
> New prompts, models, tools, integrations, or data can alter performance and revive old failures.

> **Measure Corrections, Not Just Speed**
>
> Time saved before review can disappear when errors, rewriting, privacy cleanup, or rework are counted.

Matched pairs

Each item on the left goes with the one on the right.

Task fit → How well a tool matches a defined job

Exit option → A way to export data or stop using the service

Template variable → A field replaced for each run

Stable instruction → A rule intended to remain across runs

Template owner → The person responsible for review and updates

Evaluation case → An input with expected properties used for testing

Stop-ship failure → A defect serious enough to block release

Regression case → An example retained to prevent a known failure from returning

Bounded pilot → A limited trial with success and stop criteria

Workflow handoff → A deliberate move to the pack that owns deeper scope

AI for Small Business → Team workflows, governance, and measured business pilots

Self-test

Answer each question from memory, then check your work against the answer key.

1. What should lead tool selection?

- A The longest feature page
- B A defined task and test set
- C A mascot
- D A viral demo

2. Which feature affects trust boundaries?

- A A playful name
- B Button color
- C Data retention controls
- D Brand popularity

3. What belongs in a reusable template?

- A Named input placeholders
- B Unexplained copied chat
- C A hidden final answer
- D Private data by default

4. Why version a template?

- A Hide the owner
- B Avoid all testing
- C Guarantee one perfect output
- D Track changes and regressions

5. What should an evaluation set include?

- A Normal and edge cases
- B No expected behavior
- C Only the easiest success
- D Promotional examples

Go deeper

WATCH Count the Corrections 0:26
Does artificial intelligence really save you time? Measure corrections, not just speed.

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/build-your-personal-ai-workflow

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 11 OF 14

A Map of AI

Narrow Systems, AGI, and Superintelligence · Training and Inference · Six Milestones in AI

WHAT YOU'LL BE ABLE TO DO

- Distinguish present AI systems, hypothetical intelligence categories, and the training and inference phases.

Narrow Systems, AGI, and Superintelligence

"Artificial intelligence" is an umbrella, not one machine or method. It includes systems that rank search results, detect fraud, recognize speech, recommend products, generate images, predict protein structures, and converse in natural language. These systems can be remarkably capable without sharing a single architecture or level of autonomy.

The safest starting point is to separate **observed systems** from **hypothetical categories**:

- A present-day AI system has a particular design, training process, interface, and deployment boundary. People often call such systems **narrow AI**, even when a modern model can perform many tasks.
- **Artificial general intelligence (AGI)** usually means flexible, dependable performance across a very wide range of intellectual work. There is no universally accepted test, so claims that a system "is AGI" depend heavily on the speaker's definition.
- **Artificial superintelligence (ASI)** means capability far beyond the best human performance across most important cognitive domains. It is a speculative category, not a product category that has been demonstrated.

This vocabulary is imperfect. A general-purpose language model may summarize, code, translate, tutor, and analyze images, so "narrow" can sound too restrictive. Yet breadth in a chat window does not by itself establish robust transfer, real-world autonomy, self-awareness, or reliability. Treat AGI as a disputed threshold, not as a synonym for "impressive chatbot."

The practical habit is simple: ask what a system can do **under what conditions**, how often it fails, and what evidence supports the claim. Labels come after the test.

Training and Inference

Modern machine learning is easier to understand when split into two phases.

During **training**, an optimization process changes a model's internal parameters. The system produces outputs, receives a learning signal, and adjusts its weights over many examples. Training can include pretraining on large datasets, supervised examples, preference feedback, fine-tuning, and safety or

domain-specific stages. It is usually the expensive phase, although the exact cost varies enormously by system.

During **inference**, a trained model receives a new input and produces an output. Asking for a summary, generating an image, or classifying a support ticket are inference. The prompt and conversation history can change what the model attends to in that interaction, but they do not ordinarily rewrite the base model's weights.

That distinction prevents a common misunderstanding: a conversation may become part of the current context, and a provider may separately retain data for later improvement under its policies, but neither fact means each message instantly retrains the deployed model.

The analogy "training is studying; inference is taking the test" is useful but incomplete. A model does not study like a person, and a prompt can supply temporary context the model never saw during training. For the machinery—tokens, embeddings, attention, transformers, fine-tuning, retrieval, and alignment—continue in **LLM 101**.

Six Milestones in AI

The field name dates to the **1956 Dartmouth Summer Research Project on Artificial Intelligence**. Six milestones provide a useful—not exhaustive—timeline:

1. **1956:** the Dartmouth project helped establish artificial intelligence as a field name.
2. **1997:** IBM Deep Blue defeated world chess champion Garry Kasparov.
3. **2012:** AlexNet's ImageNet result became a landmark for deep learning in computer vision.
4. **2016:** AlphaGo defeated Lee Sedol in a five-game match.
5. **2017:** "Attention Is All You Need" introduced the Transformer architecture.
6. **2022:** ChatGPT brought a conversational interface for generative AI to a mass audience.

Progress was not smooth. Periods when expectations outran results and funding fell are often called **AI winters**. The current generative wave is historically important, but it is not the beginning of AI.

Humanlike language also creates a projection trap: **anthropomorphism**, or attributing human traits and inner experience to a nonhuman system. Saying a model "knows" or "wants" something can be convenient shorthand for describing behavior. It is not evidence by itself that the model has human understanding, motives, or consciousness.

You do not need to settle philosophical questions before using AI well. Keep three layers separate:

- what the system **outputs**;
- what mechanism can explain that behavior;
- what, if anything, the behavior implies about inner experience.

The first layer is observable. The next two require additional evidence and remain active areas of research and debate.

Sources

- Dartmouth — Artificial Intelligence coined at Dartmouth
- NIST — AI Risk Management Framework
- FlashFeed, **LLM 101** pack and study guides

Key terms

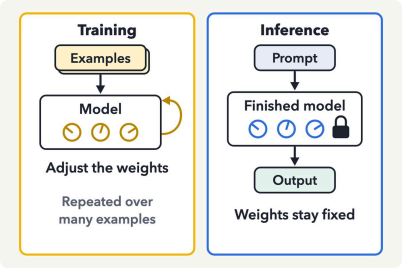
AGI

A proposed threshold for broad, flexible performance across many intellectual tasks; no universal test exists.

Inference

Using a trained model's learned values to produce an output for a request.

Key facts



FlashFeed original diagram · CC-BY-4.0

Training Changes the Model

During training, an optimization process repeatedly compares model outputs with a learning signal and adjusts internal weights. Training may use labeled examples, self-supervised objectives, feedback, or several stages.

> **Inference Uses the Model**
>
> Inference is a trained model producing an output from a new input. A prompt can change what the model considers in that interaction without automatically rewriting its underlying weights.

By the numbers

Dartmouth Summer Research Project that named the AI field

1956 year

IBM Deep Blue defeated chess champion Garry Kasparov

1997 year

AlexNet's landmark ImageNet competition victory

2012 year

AlphaGo defeated Lee Sedol in a five-game match

2016 year

Publication of “Attention Is All You Need”
2017 year

Public launch of ChatGPT
2022 year

Matched pairs

Each item on the left goes with the one on the right.

Broad fluency → Evidence of capability, not proof of an agreed AGI threshold

Inference → Using learned weights to answer without retraining them

AI winter → A period of reduced AI funding and enthusiasm

Step by step

Classify an AI activity as training or inference

1. Ask whether the model's learned weights are being changed
2. If examples or feedback change weights, classify the activity as training
3. If fixed weights process a new request, classify the activity as inference
4. Separate retrieval or tool calls from changes to the model itself
5. State the evidence for the classification and any uncertainty

A deployed system may retrieve data or call tools during inference without retraining its underlying model.

Self-test

Answer each question from memory, then check your work against the answer key.

1. Why is a superhuman Go program still a bounded system?

- A It can defeat experts within the game
- B It uses neural networks to evaluate positions
- C It searches many possible move sequences
- D Its skill does not transfer freely

2. Which status best fits superintelligence today?

- A A regulatory category for high-risk software
- B A speculative future category
- C A product term for general-purpose assistants
- D A benchmark label for any model that beats one expert

3. In the basic two-phase model, when are learned weights adjusted?

- A During training
- B During tokenization before inference
- C During decoding after each generated token
- D During inference for each prompt

4. A user asks a deployed model to summarize a memo. What phase is this?

A. Fine-tuning
B. Pretraining
C. Inference
D. Evaluation

5. What does an "AI winter" describe?

A. A temporary decline in deployed-model accuracy
B. A pause while new training data is collected
C. Reduced funding and enthusiasm
D. A period of slower chip production

Go deeper

WATCH Asking Is Not Training 0:31
Ask an artificial intelligence model for a summary. Does that prompt instantly retrain the model?

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/a-map-of-ai

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 12 OF 14

Words You'll See Everywhere

Models, Data, and Parameters · Hallucination, Bias, and Grounding · Six Key AI Concepts

WHAT YOU'LL BE ABLE TO DO

- Use collection-wide AI vocabulary to distinguish components, failure modes, interfaces, and agents.

Models, Data, and Parameters

A **machine-learning model** is a mathematical system whose behavior is shaped from examples. Instead of a programmer writing a separate rule for every possible input, training adjusts internal values so the system learns a mapping from inputs to outputs.

A **dataset** is a collection of examples used for training, evaluation, or both. A dataset is never a neutral window onto the whole world. Decisions about what to collect, how to label it, what to exclude, and how to measure quality all influence the result.

Parameters—often called weights—are the adjustable numerical values inside a model. Training data supplies a learning signal; optimization changes the parameters. The finished model is therefore related to its training data without being identical to a searchable archive of every source.

Parameter count is easy to headline and easy to misuse. More parameters can increase capacity, but the number alone does not tell you whether a model has better data, was trained efficiently, follows instructions, cites facts accurately, or fits your task. Architecture, data quality, training objective, evaluation, tools, and deployment controls also matter.

LLM 101 takes these ideas into technical depth. For this guide, remember the relationship:

> Data supplies examples. Training adjusts parameters. The model uses the learned result.

Hallucination, Bias, and Grounding

A **hallucination** is generated content that is unsupported or false while appearing plausible. A model may invent a study, combine details from different events, supply a nonexistent citation, or answer a question whose premise is wrong. The fluent surface can hide the missing foundation.

Hallucination is not the same as **bias**. Bias is a systematic pattern of distorted or unequal outcomes. It can enter through historical data, underrepresentation, labels, measurement choices, product objectives, or the setting in which a system is deployed. A fabricated citation may be a one-off hallucination; a hiring model that repeatedly disadvantages one demographic group presents a bias and fairness problem.

Grounding gives the model relevant evidence for a response. A system might retrieve documents, receive a database result, or be given approved source material in its context. Grounding can make answers more current and traceable. It does not eliminate error: the system may retrieve the wrong passage, misread a source, omit a caveat, or invent a link between facts.

Use a risk-matched verification habit:

- For brainstorming or reversible drafts, a quick human review may be enough.
- For factual claims, open the source and confirm it supports the sentence.
- For calculations, run an independent calculation or test.
- For medical, legal, financial, employment, or safety decisions, use qualified review and the governing process—not a chatbot as the final authority.

"Reduced" is the right promise for hallucination mitigation. "Eliminated" is not.

Six Key AI Concepts

Six concepts form a useful recognition set: **model**, **chatbot**, **AI agent**, **hallucination**, **grounding**, and **anthropomorphism**. The first three separate the learned system from the interface and the action loop. The last three help you reason about reliability and human projection.

An **LLM** is a model that generates and transforms language. It may power many products, but it is not the complete product.

A **chatbot** is an interface for a conversation. Behind the chat window, the product may combine an LLM with system instructions, safety filters, retrieval, web search, memory, account data, file tools, or other services. Two products using the same base model can behave differently because the surrounding system differs.

An **agent** adds goal-directed behavior over multiple steps. A typical loop is:

1. observe the current state;
2. plan or select a next action;
3. use an approved tool;
4. inspect the result;
5. continue, ask for help, or stop.

Agency is a spectrum. A system that drafts one reply and waits is less agentic than one that searches a database, updates a ticket, checks the change, and escalates an exception. "Agent" does not mean infallible or fully independent. More autonomy can multiply usefulness, but it can also multiply an error across several actions.

When evaluating an agent, ask what tools it can use, what data it can access, what limits constrain it, what actions require approval, how it handles failure, and who remains accountable.

Sources

- Google Cloud — What are AI hallucinations?
- NIST — AI Risk Management Framework
- Wang et al. — A Survey on Large Language Model Based Autonomous Agents

Key terms

Parameters
Learned internal numerical values adjusted during training to shape model behavior.

Hallucination
A plausible-seeming model output that is unsupported, incorrect, or invented.

AI agent
A system that pursues a goal through multiple steps, often using tools or taking actions.

Key concepts

Model

- A mathematical system whose behavior is shaped from examples.
- Its internal parameters are adjusted during training.
- At inference time, it maps a new input to an output.

Chatbot

- A product interface organized around a conversation.
- It may combine a language system with retrieval, memory, and safety filters.
- It usually waits for the user's next message rather than acting through a long goal loop.

AI agent

- A goal-directed system that can operate across several steps.
- It may keep state, select tools, inspect results, and change its next action.
- Its loop commonly follows observe, plan, act, check, and stop or continue.

Hallucination

- A reliability failure whose surface can look polished and plausible.
- It occurs when generated content is unsupported or false.
- An invented study or nonexistent citation is a classic example.

Grounding

- A reliability technique that gives generation an evidence base.
- The evidence may come from supplied documents, retrieval, or a trusted database.
- It reduces unsupported answers but still requires checking.

Anthropomorphism

- A reasoning trap that can be triggered by a humanlike interface.
- It treats convenient behavioral language as evidence of an inner experience.
- It means attributing human traits or motives to a nonhuman system.

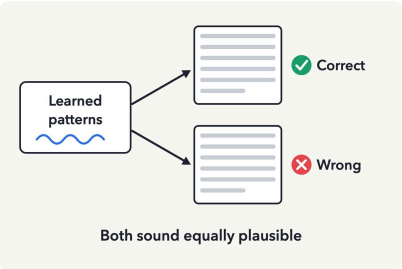
Key facts

A Model Is Learned Behavior

A machine-learning model is a mathematical system whose behavior is shaped from data. It maps inputs to outputs using learned values rather than a programmer writing every possible answer.

Data and Parameters Play Different Roles

A dataset contains examples used for learning or evaluation. Parameters are adjustable internal values. Training data influences those values, but the finished model is not simply a searchable copy of its dataset.



FlashFeed original diagram · CC-BY-4.0

Fluent Can Still Be False

A hallucination is generated content that is unsupported or false while sounding plausible. Examples include an invented citation, a fabricated quote, or an incorrect detail supplied without a reliable basis.

Bias Is Often Systematic

AI bias is a pattern of unequal or distorted outcomes tied to data, design, measurement, or deployment. Unlike a one-off fabrication, bias can repeatedly disadvantage the same group or viewpoint.

Matched pairs

Each item on the left goes with the one on the right.

Parameter count → A capacity measure, not a count of verified facts

Grounding → Relevant evidence that reduces, but cannot eliminate, hallucinations

AI agent → A system pursuing a goal across steps, tools, or actions

Self-test

Answer each question from memory, then check your work against the answer key.

1. What is the main role of a training dataset?

- A Choose the model architecture
- B Provide learning examples
- C Store every future answer verbatim
- D Set deployment permissions

2. Which description best fits model parameters?

- A Learned internal values
- B External rules attached after training
- C Raw training examples stored verbatim
- D Instructions typed by a user at inference time

3. A model supplies a plausible citation that does not exist. What happened?

- A Prompt injection
- B Data leakage
- C Hallucination
- D Bias

4. What practice gives a model relevant evidence for a response?

- A Calibration
- B Distillation
- C Quantization
- D Grounding

5. Which feature most clearly moves a chatbot toward an agent?

- A More response-formatting templates
- B A more conversational tone
- C A longer context window
- D Multi-step tool use

Go deeper

WATCH The Missing Foundation 0:29
An artificial intelligence model can write a fluent answer that cites a study. The study may not exist.

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/words-you-ll-see-everywhere

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 13 OF 14

What AI Can—and Cannot—Do

Strengths and the Reliability Gap · Energy, Water, and Physical Scale · A Regulation Snapshot

WHAT YOU'LL BE ABLE TO DO

- Choose reviewable AI uses, interpret footprint claims, and recognize jurisdiction-specific governance.

Strengths and the Reliability Gap

Current generative AI is often useful where work is **pattern-rich, well specified, reversible, and reviewable**. Common examples include:

- drafting and rewriting;
- summarizing supplied material;
- extracting fields from documents;
- classifying or clustering text;
- translating with human review;
- generating alternatives or examples;
- assisting with code, tests, and documentation.

Those are task categories, not guarantees. Performance changes with the model, prompt, source material, language, domain, and evaluation method.

The most important limit is the **reliability gap**: an output can sound better than it is. Language models are optimized to produce plausible continuations, not to make confidence in the prose equal confidence in the facts. A hesitant sentence can be correct; a decisive paragraph can be invented.

Good deployment design does not ask “Is AI smart?” It asks:

- How often is this system correct on representative examples?
- Which failures matter most?
- Can a person detect an error before harm occurs?
- Is the action reversible?
- What evidence or tool can independently check the result?
- Does the system know when to stop or escalate?

Begin with a low-consequence draft or analysis that a knowledgeable person can review. Treat high-stakes advice, rights-affecting decisions, financial transfers, and unsupervised physical actions as different risk classes—not as ordinary chat.

Energy, Water, and Physical Scale

AI feels weightless at the interface, but it runs on physical infrastructure: chips, servers, networking, storage, cooling systems, buildings, power plants, and manufacturing supply chains.

The International Energy Agency projects that **total global data-center electricity demand** could reach about **945 terawatt-hours in 2030**, a little more than double its 2024 level. AI is the largest driver of that growth, but the projection covers all data centers. Calling the entire figure "AI electricity" would erase conventional servers, storage, networking, cooling, and other digital services.

Water accounting is even more context-sensitive. Data centers may consume water directly for cooling and indirectly through electricity generation. Chip manufacturing also has a water footprint. Results vary by cooling design, climate, time of day, electricity mix, and local water stress. A single "water per prompt" number can conceal more than it explains.

Two ideas can be true together:

- data-center demand is growing quickly and can create serious local grid or water constraints;
- data centers remain one part of total global electricity demand, and efficiency, grid investment, siting, cleaner power, and workload design can change the impact.

Good footprint claims name the measurement boundary, location, period, and uncertainty.

A Regulation Snapshot

This snapshot is current as of **July 27, 2026** and is educational, not legal advice.

The **European Union AI Act** is a binding cross-sector framework organized largely by risk and use. It prohibits certain practices, imposes obligations on high-risk systems, sets rules for general-purpose AI, and requires transparency for some chatbots and generated or manipulated content. The AI Omnibus entered into force on July 27, 2026, extending key high-risk dates: certain stand-alone high-risk rules are scheduled for December 2, 2027, and rules for AI embedded in regulated products for August 2, 2028. Some transparency obligations begin August 2, 2026.

The **United States** has no single federal counterpart with the same structure. The July 2025 federal AI Action Plan emphasizes innovation, infrastructure, adoption, national security, and international leadership. Governance also comes from existing sector laws, agency actions, procurement requirements, technical standards, courts, and state laws.

The comparison is about structure, not "regulated versus unregulated." Both jurisdictions govern AI, but they assign duties through different layers and priorities. A real compliance question depends on location, sector, use, system role, risk, and effective date. Always check the current primary rule that applies to the actual deployment.

Sources

- NIST — AI Risk Management Framework
- IEA — Energy and AI

- European Commission — AI Act framework and implementation timeline
- European Commission — AI Omnibus enters into force, July 27, 2026
- White House — America's AI Action Plan

Key terms

AI governance

The rules, roles, controls, and oversight used to direct how AI is developed and used.

Key facts

> **Water Use Depends on Place and Time**
>
> Data centers may use water directly for cooling and indirectly through electricity generation and chip manufacturing. The impact varies with cooling design, climate, grid mix, and local water stress, so one global per-prompt number can mislead.

> **The U.S. Uses a More Distributed Approach**
>
> The July 2025 U.S. AI Action Plan emphasizes innovation, infrastructure, national security, and adoption. Federal agencies, sector laws, procurement rules, courts, and state laws still create additional obligations; there is no single EU-style federal AI Act.

Matched pairs

Each item on the left goes with the one on the right.

Confident tone → A writing style, not evidence that an answer is correct

AI output → A digital result produced by physical computing infrastructure

AI regulation → Rules that vary by jurisdiction, sector, and use

Step by step

Verify an important AI-generated output

1. Identify the claims or decisions that matter
2. Locate independent primary or authoritative sources
3. Check each material claim against those sources
4. Test calculations, citations, and edge cases separately
5. Record uncertainty and decide whether human approval is required

Polish and confidence are not substitutes for independent evidence.

Self-test

Answer each question from memory, then check your work against the answer key.

1. Which task is the best beginner pilot?

- A An unsupervised compliance filing
- B An automatically approved customer refund
- C A final résumé-screening rejection
- D A reviewable first draft

2. What most improves trust in an important output?

- A Repeating the same model until answers agree
- B Independent verification
- C Requesting a more detailed explanation
- D Raising the model's self-reported confidence

3. Why can one universal water-per-prompt number mislead?

- A Every facility uses the same cooling design
- B Token count is the only variable that matters
- C Systems and locations differ
- D Indirect electricity-related water is always excluded

4. What sector does the IEA's 945 TWh projection cover?

- A All data centers
- B Data centers in advanced economies only
- C AI-optimized data centers only
- D Cryptocurrency mining centers only

5. What organizes the EU AI Act's obligations?

- A Risk and use category
- B Model parameter count alone
- C The developer's headquarters country
- D Company revenue and employee count

Go deeper

WATCH Weightless on a Screen 0:28
Artificial intelligence feels weightless on a screen. But AI runs on chips, servers, cooling and power plants.

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/what-ai-can-and-cannot-do

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

LESSON 14 OF 14

Where Smart People Disagree

Will AI Take Your Job? · Catastrophic and Existential Risk · Copyright and Training Data · Timelines to AGI

> WHAT YOU'LL BE ABLE TO DO
>
> - Explain why credible estimates differ on jobs, catastrophic risk, copyright, and AGI timelines.

AI debates often fail because two people use the same word for different quantities. This lesson does not ask you to pick a camp. It teaches you to find the load-bearing assumptions.

Will AI Take Your Job?

A job is a bundle: tasks, relationships, judgment, accountability, local knowledge, and sometimes physical work. AI may automate one task, speed up another, create a new review task, and leave the rest unchanged.

That is why **exposure**, **automation**, and **job loss** are not synonyms. The IMF estimates that roughly 40% of global employment is exposed to AI, explicitly including work that AI may complement as well as substitute. Daron Acemoglu's macroeconomic model produces a much more modest near-term productivity effect because it assumes only a limited share of tasks can be profitably and reliably automated within the modeled horizon.

The optimistic case emphasizes rapid capability improvement, falling costs, new products, and complementary work that raises productivity. The cautious case emphasizes adoption friction, unreliable performance on difficult context-rich tasks, implementation costs, unequal bargaining power, and the difference between a demonstration and a production system.

When you see a dramatic jobs number, ask:

1. Is it measuring tasks, occupations, hours, wages, or people?
2. Does "affected" include assistance as well as replacement?
3. What time horizon is assumed?
4. Does the estimate model adoption and reliability, or capability alone?
5. Who gains the productivity benefit, and who bears transition costs?

Catastrophic and Existential Risk

The concerned camp argues that future systems may become much more autonomous and capable than today's models. A system able to plan, persuade, write software, use tools, and pursue long goals could create risks that are hard to contain. Because the downside could be enormous, uncertainty is itself a reason for advance testing, monitoring, access controls, security, and international coordination.

The skeptical camp starts from different evidence. Current systems are brittle, depend on human-built infrastructure, and often fail at long-horizon or real-world tasks. Skeptics warn that extrapolating to uncontrollable superintelligence may be scientifically weak and can divert attention from present harms: fraud, deepfakes, bias, surveillance, labor disruption, cyber misuse, and concentrated corporate or state power.

The 2022 Existential Risk Persuasion Tournament illustrates the gap. Median AI-domain experts assigned about a **3%** chance of AI-caused human extinction by 2100; superforecasters assigned about **0.38%**. Neither number is a measured physical constant. The groups differed in how they expected capabilities to grow, how much weight they gave base rates, and how resilient they expected institutions to be.

Long-range forecasts are difficult to score promptly. The right lesson is not "average the numbers and move on." It is to inspect scenarios, causal chains, definitions, and which evidence would change a forecast.

Copyright and Training Data

This section describes the United States as of July 2026 and is not legal advice.

"Is AI training fair use?" compresses several questions:

- Was a protected work copied while collecting or preparing data?
- Was the copy obtained lawfully?
- What was the purpose and character of the use?
- What kind of work was copied and how much?
- Could the use or resulting system harm an existing or likely market?
- Does the model memorize or reproduce protected expression?
- What jurisdiction's law applies?

In **Thomson Reuters v. Ross Intelligence**, a federal court rejected fair use where Ross copied Westlaw headnotes to build a competing legal-research product. In litigation over Anthropic, a court treated model training on books as fair use while separating that conclusion from the creation and retention of a library of pirated copies. The distinction matters: a ruling about the training purpose does not automatically excuse how every copy was acquired.

The U.S. Copyright Office's 2025 report rejected a universal answer. It explained that some training uses may be transformative, while commercial use of expressive works to produce competing content may weigh differently; licensing markets and practical access also matter. Many cases remain active, district-court opinions are fact-specific, and other countries use different doctrines.

Be suspicious of claims that one case made all training legal or illegal everywhere.

Timelines to AGI

An AGI date is meaningless until the threshold is defined. Forecasters may mean:

- a system that passes a particular benchmark suite;

- high-level machine intelligence that outperforms humans at every task;
- automation of all economically valuable work;
- a system with autonomy in the physical world;
- a company or public figure declaring that AGI exists.

These events could happen years apart.

A survey of 2,778 researchers who had published at leading AI venues produced an aggregate **50% forecast of 2047** for unaided machines outperforming humans in every possible task, assuming scientific progress continued without disruption. Individual responses varied widely. The same study's forecast for all human occupations becoming fully automatable was much later. Even within one survey, task capability and occupational automation were distinct.

Timeline optimists emphasize scaling trends, algorithmic improvement, tool use, synthetic data, automated research, and rapid capital investment. Skeptics emphasize benchmark contamination, brittle reasoning, limited real-world grounding, data and energy constraints, robotics, institutional friction, and the possibility that current methods hit hard limits.

Read every timeline as:

> Under this definition, with these assumptions, this forecaster assigns this probability by this date.

That is a forecast. "AGI arrives in 2047" is not.

Sources

- IMF — Gen-AI: Artificial Intelligence and the Future of Work
- Daron Acemoglu — The Simple Macroeconomics of AI
- Forecasting Research Institute — Existential Risk Persuasion Tournament report
- U.S. Copyright Office — Copyright and Artificial Intelligence, Part 3
- The Copyright Society — Training on Trial: Insights from Bartz and Kadrey
- Thomson Reuters v. Ross Intelligence — 2025 opinion
- Grace et al. — Thousands of AI Authors on the Future of AI

Key terms

Job exposure

The share of a job's tasks that AI could affect, assist, or automate—not a prediction that the job disappears.

Fair use

A U.S. doctrine that evaluates certain unlicensed uses through four statutory factors and case-specific facts.

Key facts

Big Numbers Often Measure Different Things

The IMF estimates about 40% of global employment is exposed to AI, including work AI may complement. Daron Acemoglu models a much smaller near-term productivity effect. These claims differ because they ask different questions.

The Skeptical Case Starts with Present Evidence

Skeptics argue that current systems remain brittle, that forecasts extrapolate too far, and that attention to speculative catastrophe can crowd out measured harms such as fraud, bias, surveillance, labor disruption, and concentrated power.

Facts and Markets Matter to Fair Use

U.S. courts have reached context-specific results. A 2025 ruling rejected fair use where Ross copied Westlaw headnotes for a competing legal tool; another found Anthropic's model training fair use but treated its pirated book library separately.

The Finish Line Moves with the Definition

Forecasts may target AGI, high-level machine intelligence, all economically valuable work, or machines outperforming humans at every task. These thresholds are not interchangeable.

Matched pairs

Each item on the left goes with the one on the right.

Job exposure → Tasks may change; it does not guarantee job elimination

AI x-risk forecast → A contested probability judgment, not a scientific consensus

Acquisition vs. use → Separate legal questions even when a later use is fair

Median forecast → The midpoint of individual probability judgments, not a deadline

Self-test

Answer each question from memory, then check your work against the answer key.

1. What is usually the better first unit for analyzing automation?

A) A firm's total headcount
B) An occupation treated as indivisible
C) An industry's annual revenue
D) A task within a job

2. Why can two credible job-impact estimates differ sharply?

- A One team publishes a longer methodology appendix
- B One report includes more industry case studies
- C Definitions and assumptions differ
- D One report uses a point estimate instead of a range

3. What drives much of the catastrophic-risk disagreement?

- A Views of future capability growth
- B Judgments about current chatbot interfaces
- C Preferences about open-source licensing
- D Views of present data-center energy use

4. In the XPT, which group gave the lower median AI-extinction forecast?

- A AI domain experts
- B Superforecasters
- C General x-risk experts
- D Non-AI domain experts

5. What is the safest general description of U.S. training law in July 2026?

- A Settled nationwide in favor of every unlicensed use
- B Governed by one rule that ignores the training corpus
- C Settled nationwide against every unlicensed use
- D Developing and fact-specific

6. What should be checked before comparing two AGI timelines?

- A Their publication locations
- B Their preferred hardware vendors
- C Their target definitions
- D Their source-code licenses

Go deeper

WATCH A Forecast, Not a Schedule 0:32
Ask researchers in artificial intelligence when unaided machines will beat humans at every task.

ONLINE This lesson on the web flashfeedlearningapp.com/study-guides/ai-for-everyone/where-smart-people-disagree

STUDY AI for Everyone in the FlashFeed app flashfeedlearningapp.com/packs/ai-for-everyone

Answer key

Lesson 1: A Good Prompt Has Parts

1. C Compare these three plans

"Compare" names a visible operation and identifies the supplied choices.

2. B A specific outcome

The request does not say whether to draft, edit, tailor, or critique.

3. D The readers are children

Age changes vocabulary, examples, and safety framing.

4. A Include details that change the answer

Useful context is selected for relevance, not sheer volume.

5. D Return a three-column table

A table with named columns is an observable structure.

Lesson 2: Learn by Showing

1. B Context and constraints

The term concerns demonstrations, not the amount of instruction.

2. D A familiar low-risk task

Simple low-risk tasks let you test the direct request cheaply.

3. C Giving one worked example

The "shot" is a demonstration, not an execution count.

4. A To separate it from live input

Boundaries clarify which text is instruction, example, and task.

5. A Varied boundary coverage

Distinct cases teach more of the decision boundary.

Lesson 3: Roles, Reasons, and Control

1. C A requested perspective

A role is a behavioral cue, not verified expertise.

2. B Review this as a copy editor

Copy editing names a concrete review lens.

3. A Show inputs and formula

Inputs and a formula can be recalculated.

4. D Only that text was produced

Length alone does not validate reasoning or evidence.

5. B Catch direction errors early

An early checkpoint prevents expensive drafting on the wrong plan.

Lesson 4: From Chatbots to Agents

1. B Multi-step action toward a goal

Agency comes from a loop that selects and executes actions.

2. A An explicit stop condition

The workflow needs success and limit rules.

3. D Read access to the chosen files

The task requires reading selected inputs, not broad authority.

4. C Sending has an external effect

A send action affects another person and merits stronger control.

5. A Sending a payment

A payment has financial consequence and may be hard to reverse.

Lesson 5: Make with Multiple Media

1. D The essential subject

The required visual content should lead the brief.

2. B Ask for more negative space

Negative space directly addresses the framing problem.

3. C 9:16 vertical

Nine-by-sixteen has the smallest width-to-height value in this set.

4. A Change over time

Motion and duration are defining dimensions of video.

5. D Generate and edit short shots

Short, planned clips make selection and continuity checks easier.

Lesson 6: Writing and Documents

1. A Meaning that must not change

Preservation rules keep fluent revision from changing substance.

2. D Ask for diagnosis before rewriting

A diagnostic pass lets the author choose which edits to accept.

3. B Audience and coverage

Selection depends on who needs what from the material.

4. C Link claims to source locations

Traceable references enable comparison with the original.

5. A Material facts and qualifications

Representation can change while substantive meaning remains fixed.

Lesson 7: Research, Learning, and Decisions

1. C Open and verify it

The source and its support for the claim both need checking.

2. B Reduce one-sided evidence

Deliberately testing alternatives helps reveal missing evidence.

3. A Wait for my attempt first

An attempt creates evidence of what the learner understands.

4. D Check it against trusted material

Generated feedback can be useful and mistaken.

5. C Unsupported precise scores

Numbers can create false precision when they lack evidence.

6. D Gives the model a clearer role you can judge

Naming the move gives the model a clearer role and helps you judge whether it was useful.

7. A Early judging can narrow the search too soon

Judging while brainstorming can narrow the search too early; generate first, then apply criteria.

Lesson 8: Data, Spreadsheets, and Code

1. C What it represents

The unit of observation determines valid aggregation.

2. D Profile and clean the data

Basic quality checks protect the analysis from structural errors.

3. B Test known small cases

Hand-computed examples show whether the formula means what you intend.

4. A Question and aggregation

The analytical question determines the appropriate grouping and visual form.

5. B Inspect context and plan

The change must fit the real repository and its constraints.

Lesson 9: Verify and Protect

1. C That it supports the claim

A source must be both genuine and aligned with the statement.

2. D Recalculate from source data

External inputs and arithmetic test the result beyond its wording.

3. B Use synthetic sample data

Synthetic examples reveal workflow needs without exposing real people.

4. A Current data-handling terms

Retention and use depend on the actual product, plan, and settings.

5. C Verify through a known channel

Voice and caller ID can be spoofed; independent contact reduces risk.

Lesson 10: Build Your Personal AI Workflow

1. B A defined task and test set

Requirements and representative cases reveal fit.

2. C Data retention controls

Retention determines how submitted information may persist.

3. A Named input placeholders

Named fields make the structure reusable and inspectable.

4. D Track changes and regressions

Versions connect prompt changes to observed performance.

5. A Normal and edge cases

Varied cases expose both routine quality and failure boundaries.

Lesson 11: A Map of AI

1. D Its skill does not transfer freely

Exceptional performance in one domain does not imply flexible competence across unrelated domains.

2. B A speculative future category

Superintelligence is a hypothetical category used in forecasting and risk debates.

3. A During training

Optimization during training changes weights; ordinary inference reads and applies them.

4. C Inference

Generating an answer from a new request is inference.

5. C Reduced funding and enthusiasm

The phrase refers to downturns after expectations outran results.

Lesson 12: Words You'll See Everywhere

1. B Provide learning examples

Examples provide the signal from which the training process shapes model behavior.

2. A Learned internal values

Parameters are numerical values adjusted during training.

3. C Hallucination

The system generated unsupported information in a credible form.

4. D Grounding

Grounding connects generation to supplied or retrieved evidence.

5. D Multi-step tool use

Planning and executing actions through tools is a central agentic capability.

Lesson 13: What AI Can—and Cannot—Do

1. D A reviewable first draft

A reversible draft creates value while keeping the consequence of an error low.

2. B Independent verification

Checks against primary evidence or tested data address correctness directly.

3. C Systems and locations differ

Model, hardware, cooling, weather, grid, and accounting boundaries all affect an estimate.

4. A All data centers

The figure is a sector-wide data-center projection for 2030.

5. A Risk and use category

The Act distinguishes prohibited, high-risk, transparency, and minimal-risk uses.

Lesson 14: Where Smart People Disagree

1. D A task within a job

Occupations combine tasks with very different potential for assistance or substitution.

2. C Definitions and assumptions differ

Time horizon, capability assumptions, exposure definitions, adoption, and task coverage can change the result.

3. A Views of future capability growth

Forecasts turn on whether progress will stay gradual or produce systems with much greater autonomy and power.

4. B Superforecasters

The 2100 medians were about 0.38% for superforecasters and 3% for AI domain experts.

5. D Developing and fact-specific

Cases turn on purpose, source, market effect, output behavior, and other facts; many suits remain active.

6. C Their target definitions

Different thresholds can yield different dates even when forecasters share similar beliefs.

Sources

Lesson sources: Google Cloud — Prompt design strategies; Min et al. — Rethinking the Role of Demonstrations; Zheng et al. — Role-Playing Does Not Improve Factual Accuracy; NIST — Generative AI Profile; OpenAI — A practical guide to building agents; NIST — Strengthening AI agent hijacking evaluations; OpenAI Academy — Image generation; Google Cloud — Configure image aspect ratio; Rombach et al. — High-Resolution Image Synthesis with Latent Diffusion Models; Radford et al. — Robust Speech Recognition via Large-Scale Weak Supervision; FTC — Approaches to address AI-enabled voice cloning; Liu et al. — Lost in the Middle; UNESCO — Guidance for generative AI in education and research; Lian et al. — A.S.E. security benchmark for AI-generated code; Office of the Privacy Commissioner of Canada — AI chatbots; FTC — Scammers use AI to enhance family-emergency schemes; NIST — AI Risk Management Framework Playbook; Dartmouth — Artificial Intelligence coined at Dartmouth; NIST — AI Risk Management Framework; Google Cloud — What are AI hallucinations?; Wang et al. — Survey of LLM-based agents; IEA — Energy and AI; European Commission — AI Act framework; White House — America's AI Action Plan; IMF — AI and the Future of Work; Acemoglu — The Simple Macroeconomics of AI; Forecasting Research Institute — XPT report; U.S. Copyright Office — Generative AI Training; Grace et al. — Thousands of AI Authors

- Google Cloud — Prompt design strategies — https://cloud.google.com/vertex-ai/generative-ai/docs/learn/prompts/prompt-design-strategies
- Min et al. — Rethinking the Role of Demonstrations — https://arxiv.org/abs/2202.12837
- Zheng et al. — Role-Playing Does Not Improve Factual Accuracy; Zheng et al. — Personas in System Prompts Do Not Improve Performances of Large Language Models (Findings of EMNLP 2024) — https://aclanthology.org/2024.findings-emnlp.888/
- NIST — Generative AI Profile — https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence
- OpenAI — A practical guide to building agents — https://openai.com/business/guides-and-resources/a-practical-guide-to-building-ai-agents/
- NIST — Strengthening AI agent hijacking evaluations — https://www.nist.gov/news-events/news/2025/01/technical-blog-strengthening-ai-agent-hijacking-evaluations
- OpenAI Academy — Image generation — https://openai.com/academy/image-generation/
- Google Cloud — Configure image aspect ratio — https://cloud.google.com/vertex-ai/generative-ai/docs/image/configure-aspect-ratio
- Rombach et al. — High-Resolution Image Synthesis with Latent Diffusion Models — https://arxiv.org/abs/2112.10752
- Radford et al. — Robust Speech Recognition via Large-Scale Weak Supervision — https://arxiv.org/abs/2212.04356
- FTC — Approaches to address AI-enabled voice cloning — https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/04/approaches-address-ai-enabled-voice-cloning
- Liu et al. — Lost in the Middle — https://aclanthology.org/2024.tacl-1.9/
- UNESCO — Guidance for generative AI in education and research — https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research
- Lian et al. — A.S.E. security benchmark for AI-generated code — https://aclanthology.org/2026.findings-acl.1569/

- Office of the Privacy Commissioner of Canada — AI chatbots — https://www.priv.gc.ca/en/privacy-topics/technology/artificial-intelligence/ai-chatbots_ind/
- FTC — Scammers use AI to enhance family-emergency schemes — https://consumer.ftc.gov/consumer-alerts/2023/03/scammers-use-ai-enhance-their-family-emergency-schemes
- NIST — AI Risk Management Framework Playbook — https://airc.nist.gov/airmf-resources/playbook/
- Dartmouth — Artificial Intelligence coined at Dartmouth — https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth
- NIST — AI Risk Management Framework — https://www.nist.gov/itl/ai-risk-management-framework
- Google Cloud — What are AI hallucinations? — https://cloud.google.com/discover/what-are-ai-hallucinations
- Wang et al. — Survey of LLM-based agents — https://arxiv.org/abs/2309.07864
- IEA — Energy and AI — https://www.iea.org/reports/energy-and-ai
- European Commission — AI Act framework — https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
- White House — America's AI Action Plan
- IMF — AI and the Future of Work — https://www.elibrary.imf.org/view/journals/006/2024/001/article-A001-en.xml
- Acemoglu — The Simple Macroeconomics of AI — https://economics.mit.edu/sites/default/files/2024-04/The%20Simple%20Macroeconomics%20of%20AI.pdf
- Forecasting Research Institute — XPT report — https://forecastingresearch.org/s/XPT.pdf
- U.S. Copyright Office — Generative AI Training — https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf
- The Copyright Society — Training on Trial
- Grace et al. — Thousands of AI Authors — https://arxiv.org/abs/2401.02843
- Google Cloud — Generate videos with Veo — https://cloud.google.com/vertex-ai/generative-ai/docs/video/generate-videos
- Google Cloud — Speaker diarization — https://docs.cloud.google.com/speech-to-text/docs/multiple-voices
- Wikipedia — Sensitivity analysis — https://en.wikipedia.org/wiki/Sensitivity_analysis
- NIST — Secure Software Development Framework — https://csrc.nist.gov/projects/ssdf
- NIST — Generative AI Profile — https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf
- IBM — Deep Blue — https://www.ibm.com/history/deep-blue
- Krizhevsky, Sutskever, and Hinton — ImageNet Classification — https://proceedings.neurips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks
- Google DeepMind — AlphaGo — https://deepmind.google/research/breakthroughs/alphago/
- Vaswani et al. — Attention Is All You Need — https://arxiv.org/abs/1706.03762
- OpenAI — Introducing ChatGPT — https://openai.com/index/chatgpt/
- IEA — Energy and AI (2025) — https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai
- The Copyright Society — Training on Trial (2026)

Image credits

Each picture in this guide is credited below, with its creator, source and license where there is one.

- **Lesson 1, Name the Deliverable:** AI-generated (gpt-image-1.5)
- **Lesson 1, Relevant Beats Exhaustive:** AI-generated (gpt-image-1.5)
- **Lesson 1, Self-Checks Still Need Review:** AI-generated (gpt-image-1.5)
- **Lesson 2, Choose a Representative Case:** AI-generated (gpt-image-1.5)
- **Lesson 2, Examples Communicate More Than Labels:** AI-generated (gpt-image-1.5)
- **Lesson 3, Task Details Usually Matter More:** AI-generated (gpt-image-1.5)
- **Lesson 3, Match the Reason to the Task:** AI-generated (gpt-image-1.5)
- **Lesson 3, Approval Can Sit Between Stages:** Wikimedia Commons · Supersonicman13 · CC BY-SA 4.0 — https://commons.wikimedia.org/wiki/File:Gepr%C3%BCfte_Servicequalit%C3%A4t.jpg
- **Lesson 3, Unknown Is a Valid Output:** Wikimedia Commons · Kwameghana(Bright Kwame Ayisi) · CC BY-SA 4.0 — https://commons.wikimedia.org/wiki/File:Foggy_landscape.jpg
- **Lesson 3, Confidence Needs Calibration:** AI-generated (gpt-image-1.5)
- **Lesson 4, Agents Operate Across Steps:** FlashFeed original diagram · CC-BY-4.0
- **Lesson 4, Reviewers Need Evidence:** AI-generated (gpt-image-1.5)
- **Lesson 6, Long Context Is Not Perfect Recall:** AI-generated (gpt-image-1.5)
- **Lesson 6, Audience Defines Plain Language:** AI-generated (gpt-image-1.5)
- **Lesson 6, Fluency Can Hide Semantic Drift:** AI-generated (gpt-image-1.5)
- **Lesson 6, Schemas Make Extraction Testable:** AI-generated (gpt-image-1.5)
- **Lesson 8, A Data Dictionary Prevents Guessing:** Wikimedia Commons · Christiane Wendehorst · CC BY-SA 4.0 — https://commons.wikimedia.org/wiki/File:Co-generated_data.png
- **Lesson 8, Profile Before Analysis:** Wikimedia Commons · Stanislav Rogozhin · Public domain — https://commons.wikimedia.org/wiki/File:Citibike_daily_ridership.png
- **Lesson 8, Security Needs Independent Controls:** Wikimedia Commons · Mmasontech · CC BY-SA 4.0 — https://commons.wikimedia.org/wiki/File:Channel_in_a_Box_Server.png
- **Lesson 9, Independent Evidence Matters:** Wikimedia Commons · Jay Walsh for the Wikimedia Foundation · CC BY-SA 4.0 — https://commons.wikimedia.org/wiki/File:How_WP_Works_books.JPG
- **Lesson 10, Placeholders Need Instructions:** Wikimedia Commons · Pejsek10 · CC BY-SA 3.0 — https://commons.wikimedia.org/wiki/File:Podac%C3%AD_l%C3%ADstek.jpg
- **Lesson 11, Training Changes the Model:** FlashFeed original diagram · CC-BY-4.0
- **Lesson 12, Fluent Can Still Be False:** FlashFeed original diagram · CC-BY-4.0

Useful links

FlashFeed

- This pack in the app: flashfeedlearningapp.com/packs/ai-for-everyone
- All packs: flashfeedlearningapp.com/packs
- How FlashFeed works: flashfeedlearningapp.com/how-it-works
- Free and open: flashfeedlearningapp.com/free-and-open

About FlashFeed

Learn anything, a little at a time.

FlashFeed turns big subjects into Knowledge Packs of short cards, quick games and study guides. Scroll when you have a minute. Go deeper when you have more. Free, with no ads and no account to make.

flashfeedlearningapp.com

Download on the App Store

Get it on Google Play

Follow

- Instagram — @flashfeedlearningapp · @flashfeed_history (history) · @flashfeed_spanish (Spanish)
- YouTube — @FlashFeedLearningApp
- TikTok — @flashfeed.learningapp
- FlashFeed Study Notes podcast — Spotify · Apple Podcasts · RSS

FlashFeed, by Recurxive LLC. November 2026 edition, built from AI for Everyone pack version 0.1.19. Typeset in Big Shoulders Display and Figtree (SIL Open Font License). Pictures are credited on the image credits page.