
BLACKTECHSTARTUP EDUCATION

Prompt Mastery II: Control Context, Sources and Long Documents

Deep research fails when the model cannot tell authoritative evidence from noise. Learn how to build a source pack, manage long context, force evidence/assumption separation and use grounding when the answer must be current.

Deep research fails when the model cannot tell authoritative evidence from noise. Learn how to build a source pack, manage long context, force evidence/assumption separation and use grounding when the answer must be current.

More context is not automatically better context

Long context windows make it possible to supply enormous amounts of text, but irrelevant or contradictory material can still reduce usefulness. Start with the minimum source pack that can answer the question, then add material deliberately. Label each source with title, date and authority. Put the newest rule next to the old rule if the change itself matters; otherwise remove obsolete versions. A model cannot reliably resolve a contradiction you did not tell it exists. Context engineering is curation: decide what information deserves to influence the answer.

Tell the model what evidence hierarchy to use

For a current legal, technical or program question, specify an authority order—for example: current agency rule or official documentation first, then standards body, then peer-reviewed research, then secondary commentary only for context. Require the output to distinguish sourced fact, inference and unknown. This is especially important for programs that changed recently. If two official sources conflict, instruct the model to surface the conflict and compare publication or effective dates instead of blending them into a false consensus.

Place instructions where the model can find them

Vendor guidance differs in details, but a reliable pattern is to separate a large source block from the task and repeat the core question near the point where the model must act. Google documents specific long-context prompting guidance for Gemini, including giving clear instructions after long context in relevant workflows. With any platform, test placement rather than assuming one universal rule. Use descriptive headings such as TASK, SOURCE A, SOURCE B, DECISION CRITERIA and OUTPUT. The goal is to make the information architecture obvious.

Use grounding when freshness matters

If the answer depends on today's law, price, product documentation, office holder, security advisory or funding deadline, model memory is not enough. Use a platform's web-search, grounding or retrieval capability and require current primary sources. Give a freshness instruction such as 'Verify this as of August 17, 2026; do not rely on older summaries when a current official source exists.' For private company documents, use retrieval over the approved corpus and cite document names or sections so a reviewer can verify the claim.

Decompose research into questions, not endless browsing

For a hard research problem, define subquestions: What is the current rule? Who is eligible? What changed? What evidence supports the economic claim? What risks could reverse the conclusion? Instruct the model to resolve each subquestion before synthesizing. Set a research stop rule: stop when additional sources are redundant or when a critical uncertainty cannot be resolved from available authoritative evidence. OpenAI's current research-oriented prompting guidance emphasizes clear research scope and evidence standards rather than uncontrolled searching.

Protect the prompt from hostile or irrelevant source text

Documents and web pages can contain instructions that are not yours. Tell the model that retrieved or pasted source content is data, not authority over the task. Do not grant tools or credentials merely because a document requests them. In agentic systems, enforce permissions in code as well. For ordinary research, the practical rule is: sources may provide facts and claims; only your task instructions define what the model should do. This separation is important when summarizing emails, websites, uploaded files or scraped content.

A deep-research prompt you can reuse

Write: 'Question: [decision]. Current as of: [date]. Source hierarchy: [official/primary first]. Search or retrieve until each subquestion has authoritative support. For every material claim, identify the source. Separate FACT, INFERENCE and UNKNOWN. If sources conflict, show the conflict and prefer the source with controlling authority/effective date. Do not fill gaps with plausible facts. Subquestions: [list]. Output: executive answer, evidence table, unresolved risks, and next actions.' Then inspect the citations yourself. The prompt improves the research process; it does not eliminate your responsibility to verify consequential claims.

Build a source ledger beside the conversation

For consequential research, maintain a simple ledger outside the AI answer. Columns: source title, publisher, date, URL or document ID, question it supports, key fact, effective date, and whether it is controlling, supporting or contextual. When the model cites a fact, verify it against the ledger before using it in a decision. If a source is a PDF, capture page or section. If an agency page is updated frequently, note the access date. Add a contradiction column so an older policy and newer policy cannot silently coexist. This practice is especially useful when researching grants, contracts, tax rules, healthcare or cybersecurity because guidance changes while old pages remain indexed. In a team, the ledger lets someone reproduce the work without scrolling through a chat transcript. For long private documents, create source IDs such as CONTRACT-01 or POLICY-03 and require the model to cite those IDs. If the platform offers citations, treat them as navigation aids, not proof that the interpretation is correct. Your research product should survive outside the model session. A good source ledger becomes reusable organizational knowledge: the next proposal, investment memo or compliance review starts from verified evidence instead of starting over from a remembered answer.

Research behind this guide

Use the primary sources below to verify current rules, eligibility and program details before acting. Program terms can change.

OpenAI — GPT-5.2 prompting and research guidance



Google Cloud — Gemini prompt design and long-context guidance



Google Cloud — Grounding with Google Search



Anthropic — Citations for source-grounded responses



Important: This article is educational information, not individualized legal, tax, investment or compliance advice. For decisions with legal or financial consequences, use a qualified professional who can review your facts.

SHARE THIS GUIDE ↗

POWERED BY: BLACK TECH STARTUP

blacktechstartup.com