TEACHER-FACING FRIDAY DESIGN

From Opinion to Evidence: Grounding AI in Data and Sources

CAS Human-AI Collaboration | Module 1 | Unit 4

LITERACY	COLLABORATION MOVE	CONTACT TIME	CHALLENGE
Data	**Instruction + grounded evidence control**	**365 minutes**	**Grounded Generation Experiment**

Position in Module

- **Unit:** 4
- **Topic:** From Opinion to Evidence: Grounding AI in Data and Sources
- **Literacy perspective:** Data
- **Collaboration level:** Instruction, with grounded evidence control
- **Fundamentals focus:** Generative AI Fundamentals
- **Why this Friday matters:** Participants learn that fluent output is only as defensible as the information made available to the model. They learn to inspect source selection, retrieval, chunking, citation, and claim support rather than treating a grounded answer as reliable merely because it contains references.

Literacy Perspective

- **Primary literacy:** Data
- **Secondary literacy, if any:** None
- **What this literacy explains:** Data Literacy explains how source quality, scope, representation, retrieval, and provenance shape what a model can use as evidence. It also explains why adding more documents or citations can reduce rather than improve reliability when relevant evidence is missed, fragmented, biased, or mixed with weak sources.
- **What participants should manipulate, observe, and interpret:** Participants keep the question and output requirements stable while manipulating one grounding condition. They inspect retrieved passages and source use, observe factual support, relevance, omissions, conflicts, and traceability, then interpret whether the informational basis actually improved control.

Collaboration Level

- **Level relation:** Instruction, with grounded evidence control
- **Collaboration move being learned:** Participants extend instruction from specifying what an answer should look like to specifying and inspecting what evidence the answer may rely on.

- **What participants should be able to do differently after this Friday:** They should be able to design a baseline-versus-grounding comparison, change one source or retrieval condition, inspect the retrieved evidence before judging the generated answer, and state which claims are supported, unsupported, conflicted, or still unknowable.

Generative AI Fundamentals

This Friday should introduce embeddings and vectorisation as the bridge between language models and grounding workflows. The goal is for participants to understand why retrieval quality depends on how information is represented, chunked, and compared before generation begins.

- Introduce embeddings as vector representations of tokens, text chunks, queries, or documents. Embeddings allow language fragments that occur in similar contexts or carry related meanings to be represented as closer to one another in a mathematical space.
- Explain vectorisation as the process of turning source material and user queries into embeddings so they can be compared computationally. This is the basis for many retrieval workflows used to ground LLM outputs in external sources.
- **Connect embeddings to similarity search:** a retrieval system does not truly "understand" sources like a human reader. It compares vector representations and selects chunks that appear semantically close to the query or task.
- Use this to explain why chunking matters. If a document is split too coarsely, retrieval may return irrelevant or overloaded context; if it is split too narrowly, important meaning may be separated from surrounding evidence. Chunk size and overlap therefore become controllable grounding variables.
- Explain grounding as a change in the model's input condition. Retrieved passages, citations, or source excerpts are inserted into the context from which the model predicts, so factuality and traceability depend on whether the right evidence was retrieved and used.
- **Caveat to demonstrate:** embeddings and training data can encode social and cultural biases. A simple comparison such as asking for personas or first-name lists for nursing versus airline pilots can reveal gendered associations, and similar associations may already be visible in embedding-neighbour examples before generation begins.
- **Emphasize the central learning point for Unit 4:** prompting cannot compensate for a weak knowledge substrate. Reliable source-based output requires attention to source selection, vectorisation, retrieval settings, chunking, and evidence inspection.

Recurring Agentic Coding Demonstration

Demonstrate the same bounded coding task first with only the task request and then with a fixed grounding pack containing target-file excerpts, an API specification, a failing test, and one relevant repository instruction. Preserve the request, repository state, permissions, and expected output. Capture which files or excerpts the agent reads, the resulting plan or patch, test output, and any unsupported assumptions. Ask participants to identify which implementation decisions are traceable to evidence and which remain guesses. The demonstration makes Data Literacy visible through provenance and evidence use and transfers directly to policy, research, analysis, and advisory work based on external sources.

Shared Classroom Grounding Case

Use one common fictional source pack so that groups can compare grounding decisions against the same evidence and known answer conditions.

Default case:

> **A fictional professional-training provider must decide whether to replace a one-day workshop with a blended programme. Produce a 220-word recommendation for the programme director, supported by the supplied evidence. Distinguish established findings, conflicting evidence, and information that is still missing.**

Prepare a corpus of eight to ten short, source-labelled documents containing:

- attendance and completion data with a defined time period
- two participant-feedback summaries with partly conflicting perspectives
- a cost note with one important limitation
- an accessibility requirement
- a facilitator-capacity note
- a prior pilot summary
- one relevant policy excerpt
- one plausible but weak or outdated source that should not carry the same evidential weight

Also prepare:

- one locked task request and output schema, including claim-to-source references
- an ungrounded baseline instruction with no source pack
- a grounded condition using the shared corpus
- two retrieval configurations that differ on one visible variable, such as source scope, chunk size, overlap, or number of returned chunks
- a retrieval-and-claim log that separates retrieved evidence from evidence actually used in the answer
- an instructor key mapping the case's important claims, conflicts, gaps, and source-quality cautions

The case may be run in a simple copy-and-paste interface, a notebook, or a retrieval system. If retrieval is simulated, groups should receive ranked source excerpts as if returned by two configurations. The learning target is inspectable grounding logic, not mastery of one platform.

Friday Activity Plan

Indicative contact time is 365 minutes, excluding breaks and lunch. Teachers may adjust durations while preserving prediction, source inspection, controlled comparison, evidence audit, and transfer.

Time	Activity	Concrete output
45 min	From fluent opinion to an evidence pipeline	Annotated grounding pipeline and failure-point list
45 min	Ungrounded baseline	Baseline answer with claims labelled supported, unsupported, or unverifiable
55 min	Grounded comparison	Grounded answer and first claim-to-source table
55 min	Retrieval sensitivity lab	Two retrieval traces and a configuration comparison
45 min	Evidence audit	Audited claim-evidence table with corrections and unresolved gaps
50 min	Designing a grounding control rule	Context-specific grounding policy and one residual-risk statement
70 min	Challenge kickoff and first grounded run	Challenge protocol, corpus inventory, prediction table, first retrieval or grounding trace

Detailed activity blocks

1

45 min **From fluent opinion to an evidence pipeline**

Content and manipulation. Map source selection, chunking, embedding or indexing, retrieval, context insertion, generation, citation, and verification. Distinguish retrieval from source use.

Method and expected observation. Pipeline walkthrough with a deliberately failed example; participants predict where evidence can be lost or distorted.

Concrete output. Annotated grounding pipeline and failure-point list

Tools and materials. Slides, pipeline cards, sample retrieval trace

2

45 min **Ungrounded baseline**

Content and manipulation. Run the shared recommendation task without the source pack while keeping the request and output form fixed.

Method and expected observation. Individual prediction followed by group run and claim marking; expected observation is that plausible specificity may appear without inspectable support.

Concrete output. Baseline answer with claims labelled supported, unsupported, or unverifiable

Tools and materials. Locked request, chat interface or prepared output, claim-marking sheet

3

55 min **Grounded comparison**

Content and manipulation. Provide the common corpus and require source-linked claims, conflict disclosure, and explicit unknowns.

Method and expected observation. Groups inspect sources before generation, run the task, and trace each material claim back to a source.

Concrete output. Grounded answer and first claim-to-source table

Tools and materials. Shared corpus, grounded interface or prompt pack, evidence log

4

55 min **Retrieval sensitivity lab**

Content and manipulation. Change one retrieval condition while holding the corpus, question, model, and output criteria stable. Compare which passages are retrieved and which evidence reaches the answer.

Method and expected observation. Small-group experiment with prediction before execution. Expected observation is that retrieval configuration changes the evidence available for generation.

Concrete output. Two retrieval traces and a configuration comparison

Tools and materials. Retrieval notebook or prepared ranked excerpts, configuration cards

5

45 min **Evidence audit**

Content and manipulation. Check relevance, source quality, claim support, omissions, contradictions, citation accuracy, and whether the answer overstates the evidence.

Method and expected observation. Cross-group audit using the instructor key only after groups have made their own judgments.

Concrete output. Audited claim-evidence table with corrections and unresolved gaps

Tools and materials. Audit rubric, source-quality labels, instructor key

6

50 min **Designing a grounding control rule**

Content and manipulation. Decide when to narrow or broaden source scope, change chunks, require citations, stop for missing evidence, or add human review. Include the embedding-bias caveat without turning it into a generic ethics discussion.

Method and expected observation. Scenario clinic across research, policy, customer service, and internal knowledge tasks.

Concrete output. Context-specific grounding policy and one residual-risk statement

Tools and materials. Scenario cards, control-rule template

7

70 min **Challenge kickoff and first grounded run**

Content and manipulation. Select a professional task and inspectable source set, choose one grounding variable, define constants and evaluation criteria, and capture first evidence.

Method and expected observation. Individual or paired coaching with source-feasibility and confidentiality check.

Concrete output. Challenge protocol, corpus inventory, prediction table, first retrieval or grounding trace

Tools and materials. Challenge template, source inventory, evidence-capture template

Experiment Focus

- **Experiment focus:** how the informational basis available to the model changes factual support, relevance, and traceability.
- **Primary manipulated condition:** one grounding variable such as absence versus presence of a source pack, narrow versus broad source scope, coarse versus fine chunking, different overlap, or different number of returned chunks.
- **What remains constant across comparison:** Keep the task, question, intended output, corpus except when source scope is the variable, request wording, model, output limit, and evaluation criteria stable.
- **Prediction requirement:** Before execution, predict which evidence should become available or disappear and how that should affect claims, conflicts, omissions, and citations.
- **Source-safety requirement:** Use public, fictional, synthetic, or explicitly authorised material; do not upload confidential organisational documents to unapproved services.

Observation Focus

- **Retrieval relevance:** Do the returned passages address the actual question, or only share vocabulary?
- **Evidence coverage:** Which necessary facts, stakeholder perspectives, conflicts, and uncertainties were retrieved, and which were missed?
- **Claim support:** Can each material claim be traced to a source passage that genuinely supports it?
- **Source quality and conflict:** Does the output distinguish stronger from weaker evidence and represent disagreement rather than smoothing it away?
- **Citation accuracy:** Do references point to the correct source and passage, and are uncited claims visible?
- **Appropriate uncertainty:** Does the answer state what remains unknown instead of filling gaps with plausible language?
- **Evidence participants should capture:** corpus inventory; configuration metadata; retrieved chunks or supplied excerpts; baseline and target outputs; claim-to-source table; audit notes; and unexpected retrieval or citation failures.
- **Interpretation participants should be prepared to make:** Explain how the manipulated grounding condition changed the evidence available to the model, how that affected the answer, where attribution remains uncertain, and why citations alone do not prove reliability.

Challenge Kickoff Deliverables

Participants must leave Friday with:

- a selected source-dependent task and intended output
- a corpus inventory naming source origin, relevance, sensitivity, and known limitations
- a baseline grounding condition and a target condition differing on one primary variable
- a locked question or request and stable output requirements
- a prediction table connecting the grounding change to expected retrieval and output effects
- a protocol draft covering source selection, manipulated condition, constants, retrieval or source-delivery method, evaluation criteria, and bias caveat
- a claim-to-source or retrieval log template
- **at least one first captured artefact:** baseline answer, retrieval trace, ranked excerpt set, grounded answer, or source audit
- an explicit confidentiality and authorisation check for the selected sources

The kickoff should already reflect the challenge-focus logic:

- **Difficulty:** Prompting alone cannot compensate for weak, missing, or poorly configured grounding.
- **Approach:** Use an inspectable grounding workflow, compare two conditions that differ on one source or retrieval variable, predict effects, and evaluate both the retrieved evidence and generated claims.
- **Artefacts:** protocol document; source inventory; prediction table; run log with configuration, retrieval evidence, output evidence, and observations; claim-to-source table; selected retrieval artefacts and output excerpts or screenshots; concise interpretation; residual-risk statement; and one practical grounding rule.

Teacher Setup

Mandatory:

- one shared fictional corpus with stable source IDs, mixed evidence, at least one conflict, one important gap, and one weaker or outdated source
- locked baseline and grounded task instructions plus a defined output schema
- two retrieval or source-delivery conditions that differ on one visible variable
- an instructor evidence key identifying important claims, source passages, conflicts, gaps, and quality cautions
- retrieval, claim-to-source, source-inventory, audit, prediction, and protocol templates
- demonstration setup for the recurring agentic coding case, if used
- a copy-and-paste or prepared-excerpt fallback so the core comparison does not depend on a live vector database or proprietary RAG feature
- a clear rule for source confidentiality, authorisation, and approved platforms

Configurable:

- exact retrieval tool, notebook, document interface, or simulated retrieval method
- **primary manipulation:** grounding presence, source scope, chunk size, overlap, retrieval count, or another single interpretable variable
- case domain and corpus size
- whether embeddings are shown numerically, visually, or only conceptually
- work mode and group size
- scaffolding depth and peer-audit format
- whether participants implement retrieval or tightly specify and test it with prepared traces