

user interviews



# AI in UXR Context Engineering **LAB**

Kaleb Loosbrock

Staff Researcher | Instacart

Founder AIxUXR Community

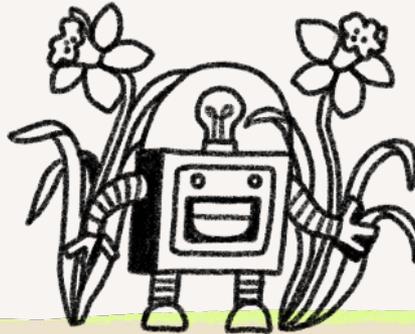


## Meet the Instructor

- IND → CHI → SF → CHI
- Doing UXR for 12+ years
- Interesting facts:
  - Product Manager for 6+ years
  - Making meaning is my superpower
  - Traveled to 37 countries & counting!
- UX Insights Conference : April 20–22nd, 2026!

**LET'S DO A QUICK POLL!**

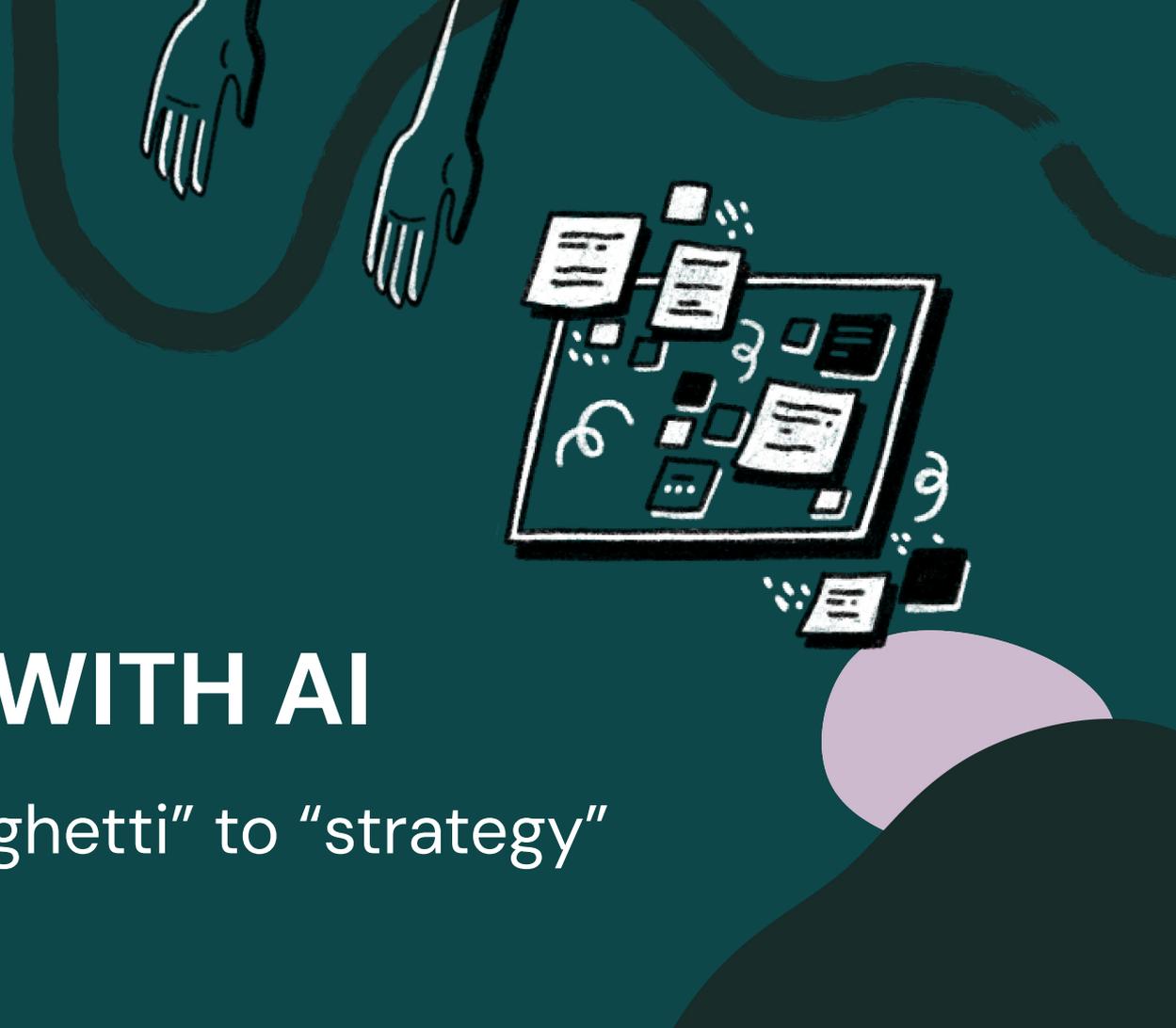
Did you take the AI in UXR 101 course?



THE LAB

# HANDS ON WITH AI

Moving from “spaghetti” to “strategy”



# AI in UXR: Context Engineering **LAB**

## Course Objectives:

- **Prompt → Context Engineering**  
Move beyond basic prompts to structure context for complex tasks (the "CRAFTe for Context").
- **Execute Practical Workflows**  
Transform a raw VTT transcript into a structured, analyzable summary for your stakeholders and post-session notes.
- **Navigate AI Parameters**  
Learn to manipulate "Temperature" and "System Instructions" to toggle between creative brainstorming and rigorous analysis.

## What You'll Get:

- ✓ **The Operator's Playbook:** Advanced techniques for hallucination mitigation.
- ✓ **Advanced Techniques** for reducing hallucinations & improving quality
- ✓ **Hands on experience** leveraging AI tools and settings to significantly improve the outcomes
- ✓ **201 Certification:** A verifiable credential for your professional profile.



**Bonus:** New User Interviews customers will get a credit for 3 free participants

# Session Breakdown

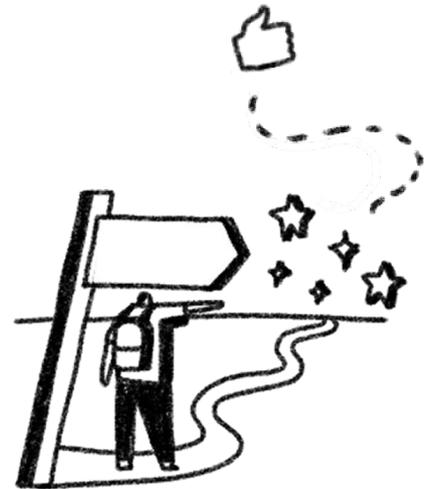
**TODAY**

## AI in UXR: Context Engineering

- The Refresh: Context Engineering Review & Updates
- The Architecture: Workflow Overview & Safety Briefing

**5 min break**

- The Demo: Live tool & workflow
- The Lab: Experimentation Time (Hands-on)
- The Debrief: Q&A



DEEP DIVE

CASE STUDY

# Walking the walk



## BACKGROUND

### Only half of the research team getting value from AI

Despite the AI-First mandate from Leadership, Instacart's Research team was struggling with realizing the prophesized gains in their day-to-day—mainly from a lack of knowledge, experience, and confidence.

**50%**

Report effectively use AI for complex research tasks

**40%**

Felt AI outputs were good enough to meet their research needs

**50%**

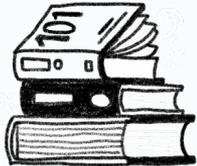
Report feeling that AI helped them complete tasks quickly

# THE SOLUTION

## We tackled it from multiple angles...

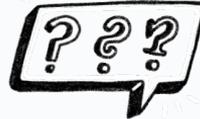
### HOSTED AI WORKSHOPS

Hosted multiple AI workshops centered on educating on the fundamentals.



### CREATED SPACE FOR CONVERSATION

Created "AI Fridays" for teams to share and learn from each other and about AI without judgement



### DEDICATED OKRS

Sanctioned and executed on 2 Dedicated Research OKRs to measure and improve AI literacy & unlocks



# The Unlock...

## TASK ANALYSIS TO BUILD AN AI-ASSISTED WORKFLOW & PLAYBOOK

Hosted a multi-day AI sprint with the Research and Ops teams to map out various tasks (lit review, discussion guide, etc.) with the goal of designing AI-assisted workflows & playbooks.

### 2 Ideate an AI-assisted workflow

#### Brainstorm!

~30-45 MIN

You've mapped, presented and discussed your individual processes and identified the pain points. Now it's time to imagine the ideal workflow with AI involved.

Sketch out your proposed solution and the prompts you might use.

What is our vision for a better way, using AI?

#### INSTRUCTIONS

- Start mapping out step-by-step the ideal workflow. Label your steps and any potential alternative flows that might be explored.
- Indicate on each step the following:
  - The Human's Action** - These are actions that a human must perform. Create new file, Create project in Dovetail, etc.
  - The Human's Oversight** - These are where you feel that the Human Researcher must be involved. Review for full coverage, Check for factual accuracy, Assess for tone
  - The AI Action** - This is the step that you think AI could perform in part of your workflow. Summarize text, Generate Images, Identify key terms
- Indicate which tools you recommend using as a part of each step (bonus points for the AI model!) Google Docs, Figma, Google Vertex, Google Pro 2.5, etc.
- Identify any potential pitfalls or risks: What can go wrong? This focuses on the quality and integrity of the output.

#### Prompts to get you started

- What parts should stay human?
- What are various outputs/inputs to various tasks and sub-tasks?
- Where and how could AI assist?
- What would the ideal AI-human collaboration look like for this task?
- How does your task need to adapt and evolve to better include your AI partner?
- Are there a set of best practices that you may or may not be adhering to (no judgement!) Maybe AI can help you better adhere to it?
- If we offload a step to AI, what could potentially go wrong?

#### Potential Workflows for Generating Summaries with AI



## THE RESULTS

Researchers were more comfortable, confident, and effective at using AI.



**90%**

Report effectively use AI for *complex* research tasks

**2/3**

Now, felt AI outputs were good enough to meet their needs

**33%**

Reduction in task time\*

\*Reduction figure is specific to today's workflow task: generating post-session notes & summaries.

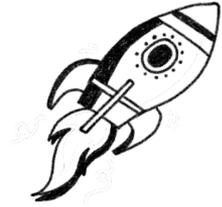
# What's up next

1

**Context Engineering 101**

2

**Workflow & Tool Walkthrough**



3

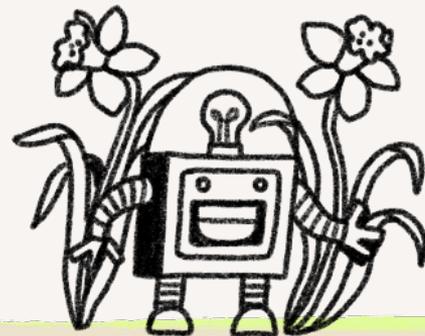
**Hands On Lab!**



## LET'S DO A QUICK POLL!

**Which statement best captures how you currently operate with AI for generating research outputs?**

- A. The Passive Reviewer:** I use AI for baseline tasks (transcription/summaries) but treat the output as a draft that requires manual correction.
- B. The Socratic Operator:** I use AI as a sparring partner to challenge my own bias, asking the model for alternative interpretations or counter-arguments.
- C. The Context Engineer:** I architect the input (data chunking, few-shot examples, and constraints) to strictly steer the model toward deterministic, high-fidelity results.
- D. The Research Orchestrator:** I design and deploy custom agents or multi-step systems to automate complex synthesis across entire datasets for my team.



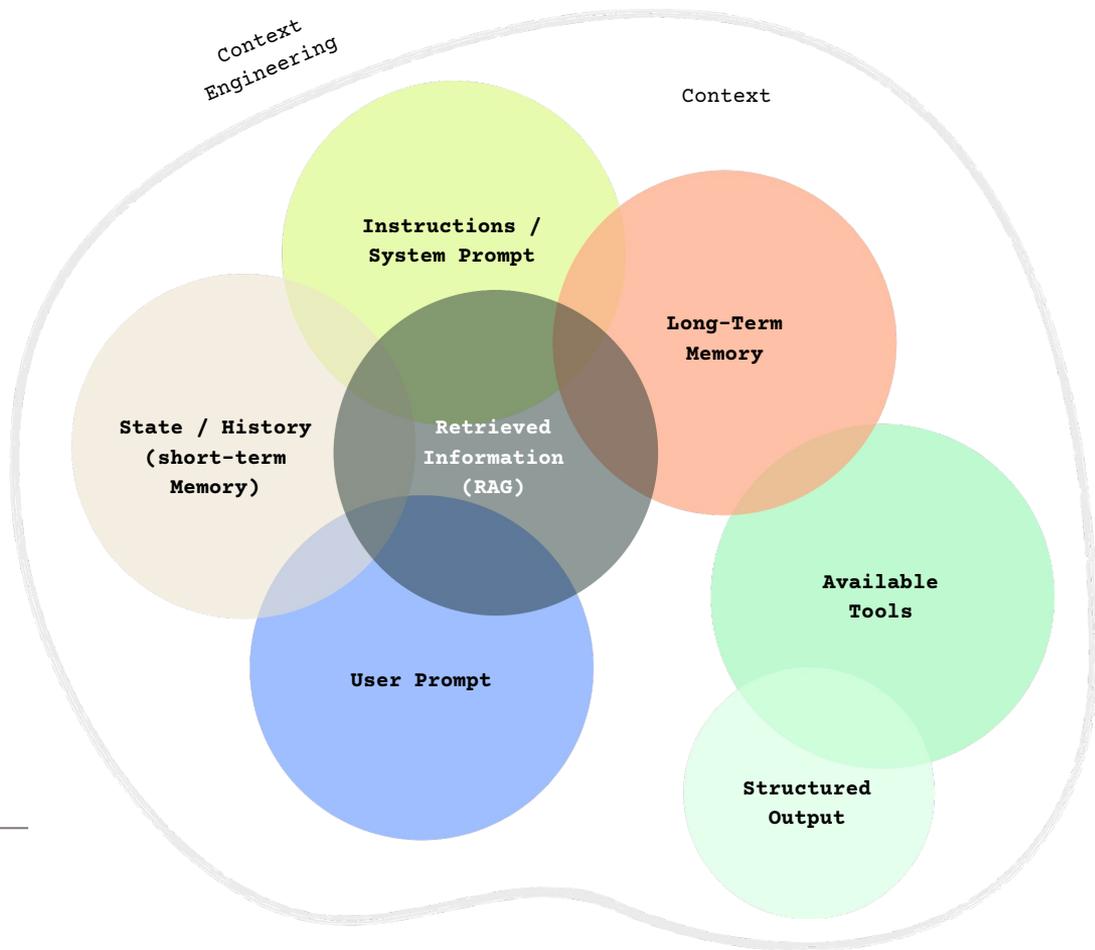
# Subtext of Context

Intro Techniques for Effective Context Engineering



# Context Engineering

While prompt engineering centers on crafting precise instructions, context engineering concentrates on architecting the entire world of knowledge that the model needs to accurately interpret instructions and act effectively.



# Why do context engineering?

## IMPROVED QUALITY

A context-architecture guide for analytics agents claims that adding structured “five levels of context” can move systems from roughly 20% baseline accuracy to 94%+. <sup>1</sup>

1. [Promethium \(2025\) Context Architecture for AI Analytics: The Five Levels That Determine Accuracy](#)

## REDUCED TIME & COSTS



Sprinklr case studies report up to 50% accuracy improvements and 60% lower compute costs through targeted context optimization. <sup>2</sup>

2. [Sprinkler Blog\(2025\).The Context Engineering Advantage: How Leading Enterprises Scale AI Agents Successfully](#)

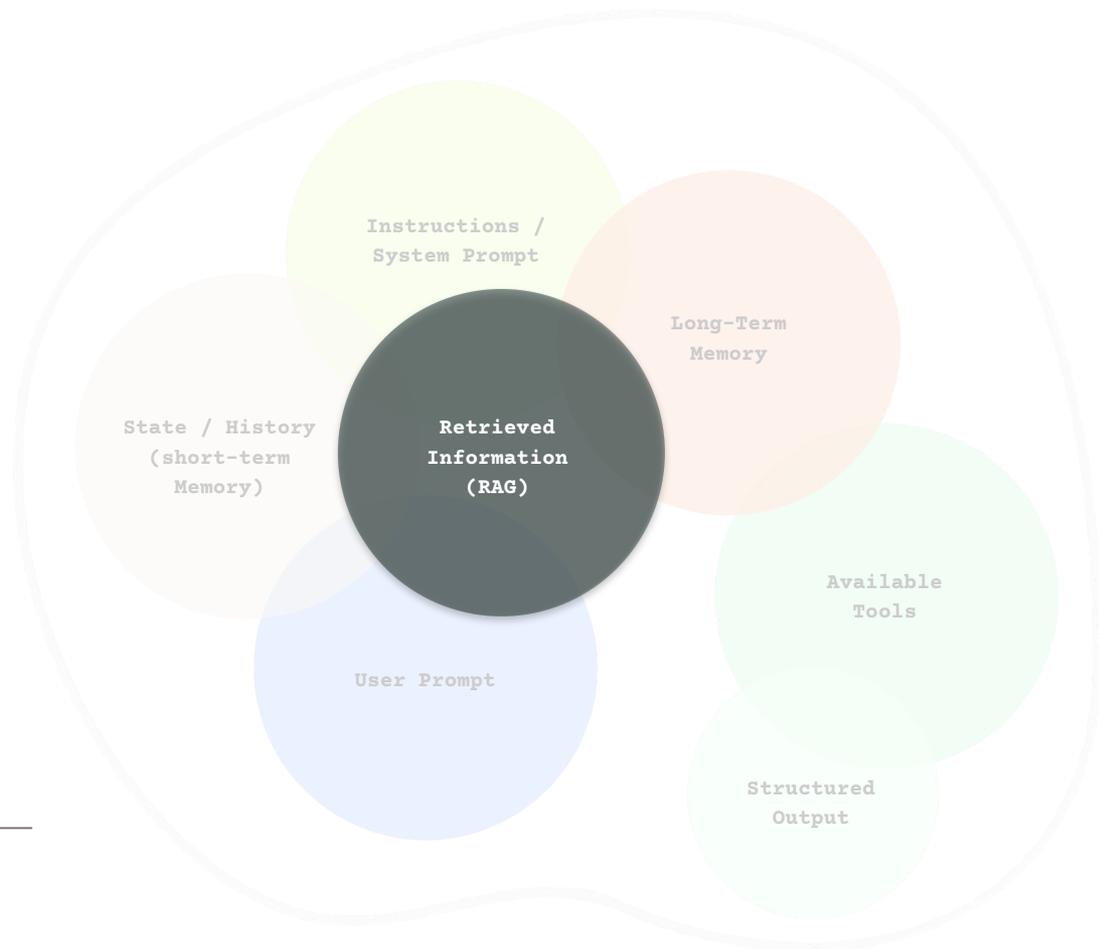
## LESS HEADACHES



Hopefully at the end of the day, that means less headaches for you, your team, stakeholders, and partners.

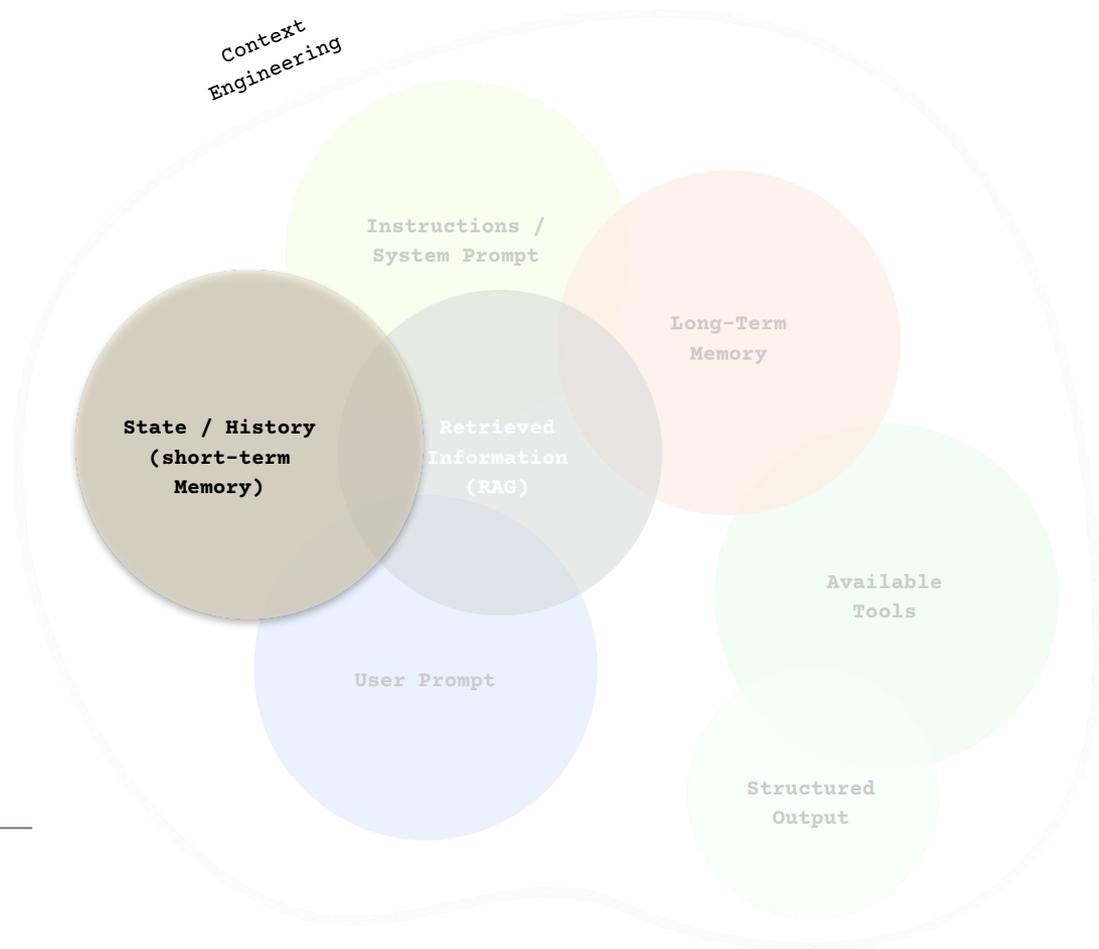
# Retrieved Information

Retrieval Augmented Generation (RAG) is a technique that retrieves information from curated “knowledge base (e.g. vector databases) to ground the prompt and improve accuracy, relevancy, and reliability.



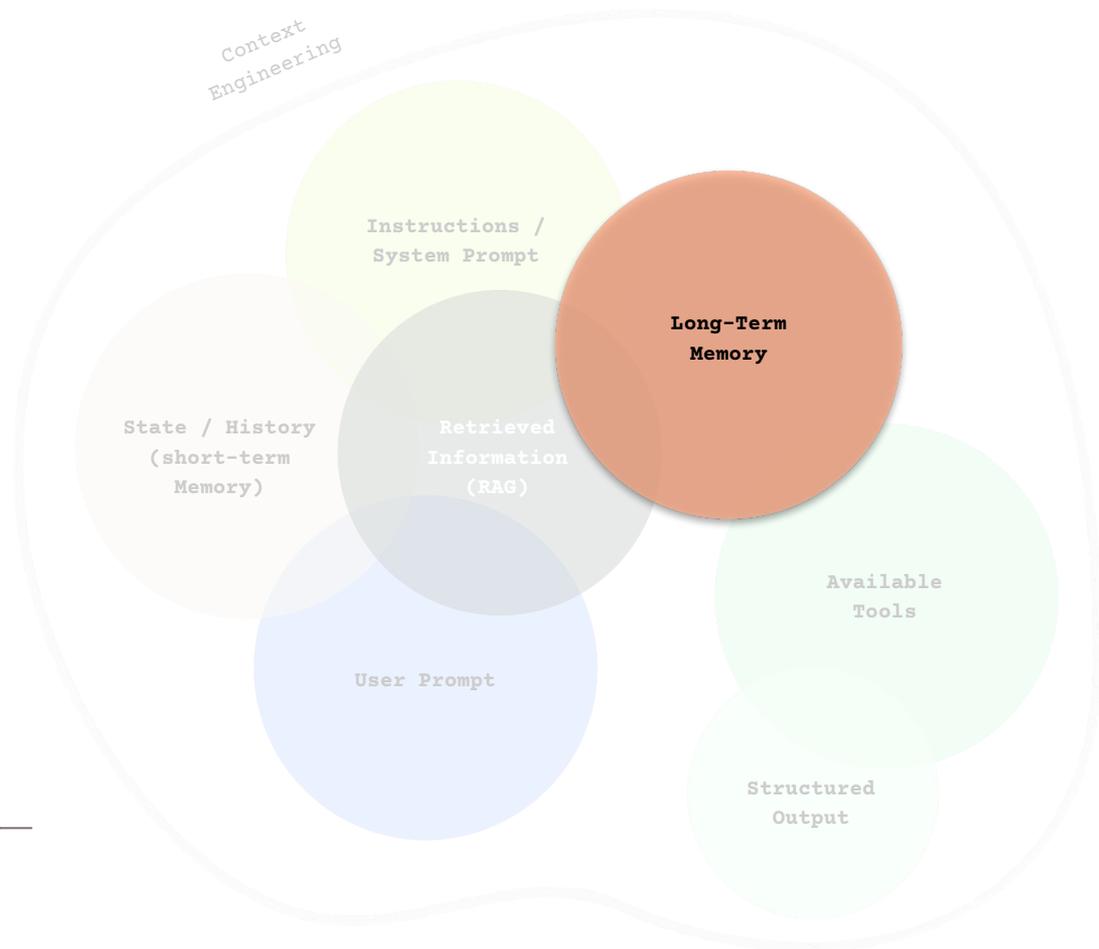
# State/History

Acts as the AI's short-term memory by keeping track of everything said during a single session to ensure the conversation stays consistent and on track.



## Long-term memory

AI's ability to remember who you are and how you like to work even after you close the chat window and come back weeks later. It acts as a permanent "filing cabinet" for your projects and conversations.



# TOKENS = MEMORY

## Context Window

Short-term memory or attention span



## Input Tokens

The tokens as part of your prompt



## Output Tokens

The tokens used as part of the models' response

**Pro-tip 1:** For analysis, utilize models with higher token limits.

**Pro-tip 2:** Checkout [OpenAI's Tokenizer](#) to see just how text translates to tokens!

Tokens

31,718

Characters

84274

WEBVTT

00:03:49.120 --> 00:03:50.400

<v Moderator>Hola, buenas tardes.

00:03:51.710 --> 00:03:52.430

<v P12 (Marco)>Buenas tardes.

00:03:52.750 --> 00:03:53.710

<v Moderator>¿Cómo estás?

00:03:54.510 --> 00:03:55.150

<v P12 (Marco)>Muy bien.

00:03:55.390 --> 00:03:56.670

<v Moderator>Me alegra, Marco.

00:03:56.670 --> 00:03:58.350

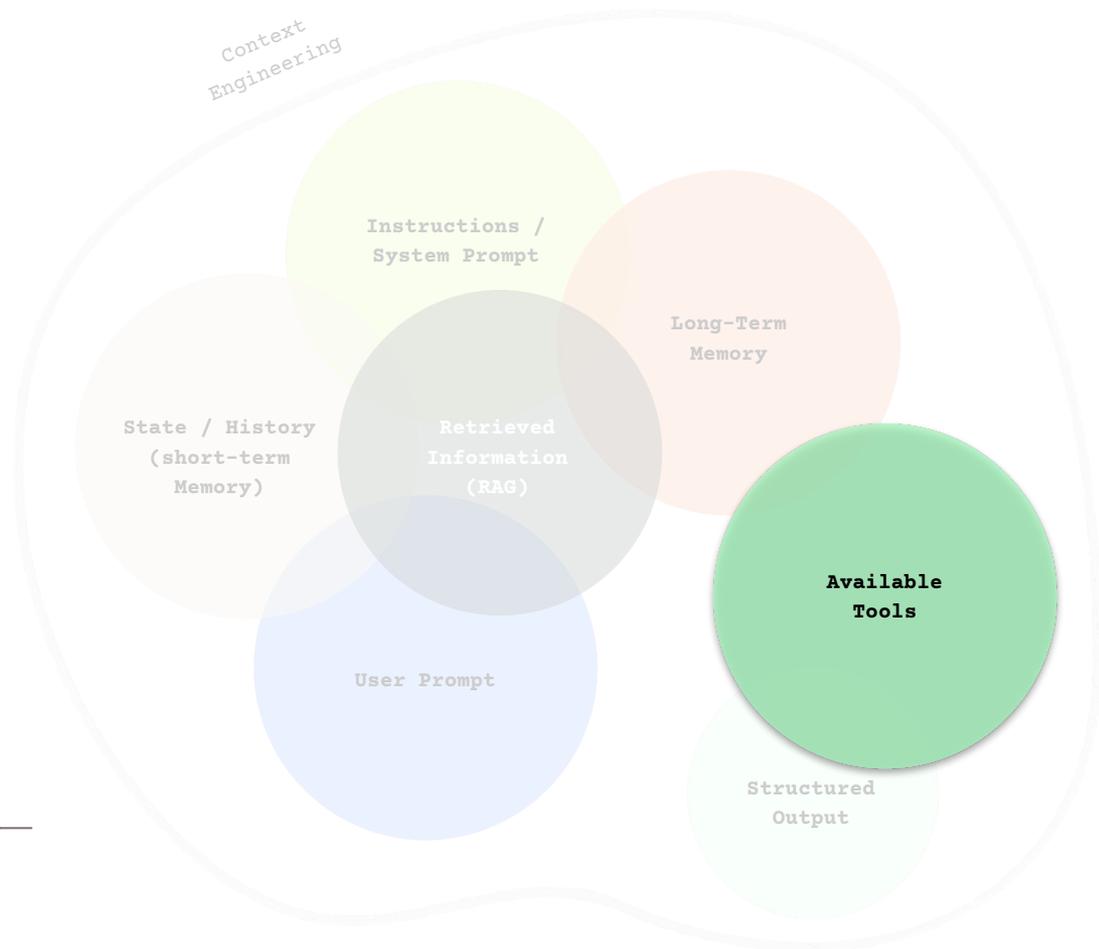
<v Moderator>Mi nombre es Marlana. Un placer.

Text

Token IDs

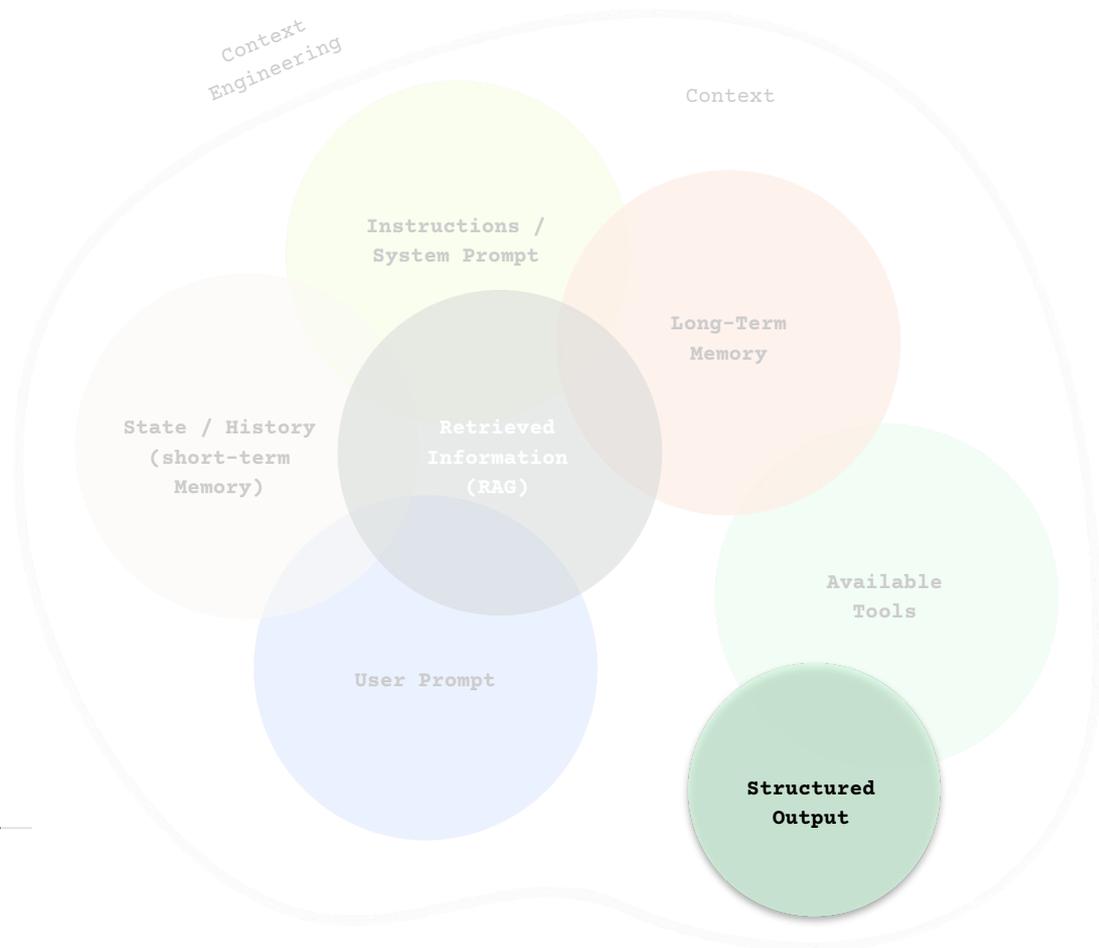
## Available Tools

AI's "Utility Belt". They are specialized mini-programs that the AI can "call" to do things it can't do with just words, like searching the live internet, running math equations or code, or making images.



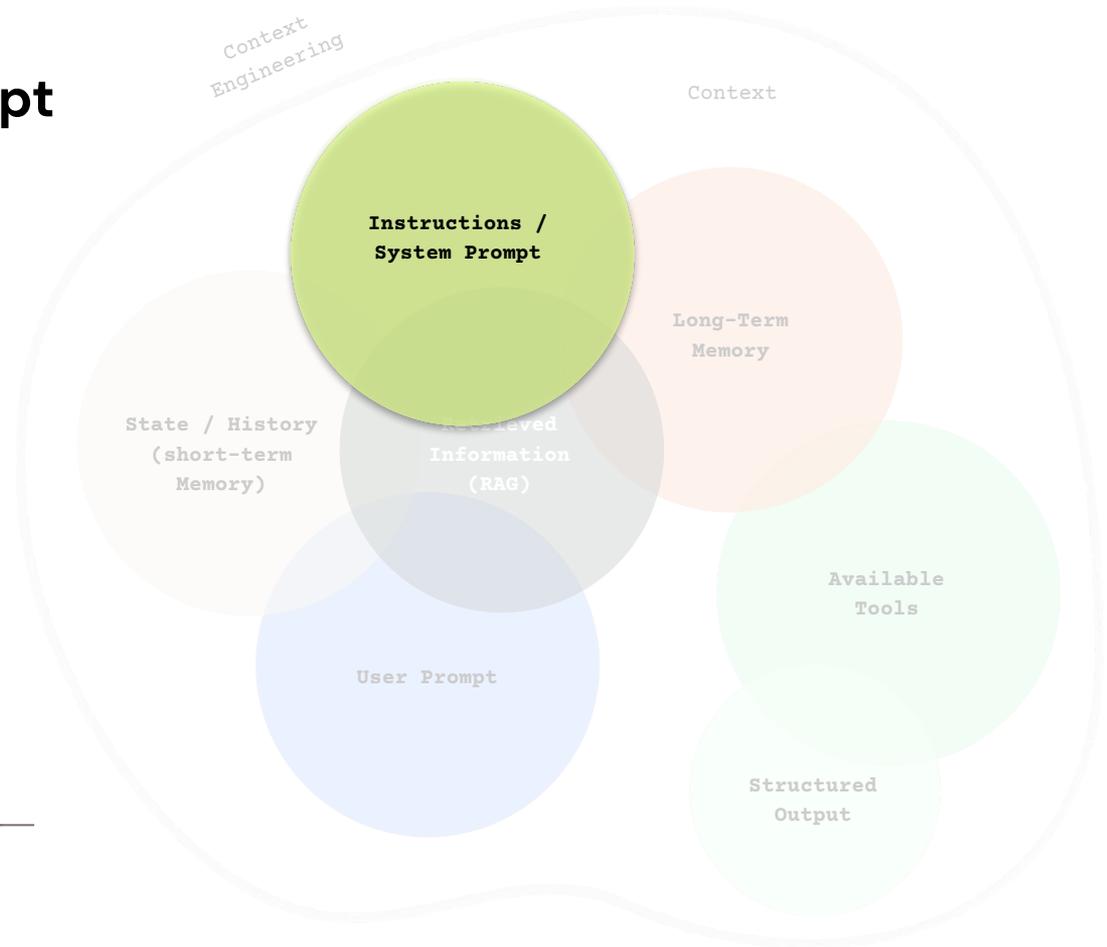
# Structured Output

The AI's ability to organize its answer into a very specific format (like a table, a list, or XML tags) instead of just writing a long, messy paragraph.



# Instructions/System Prompt

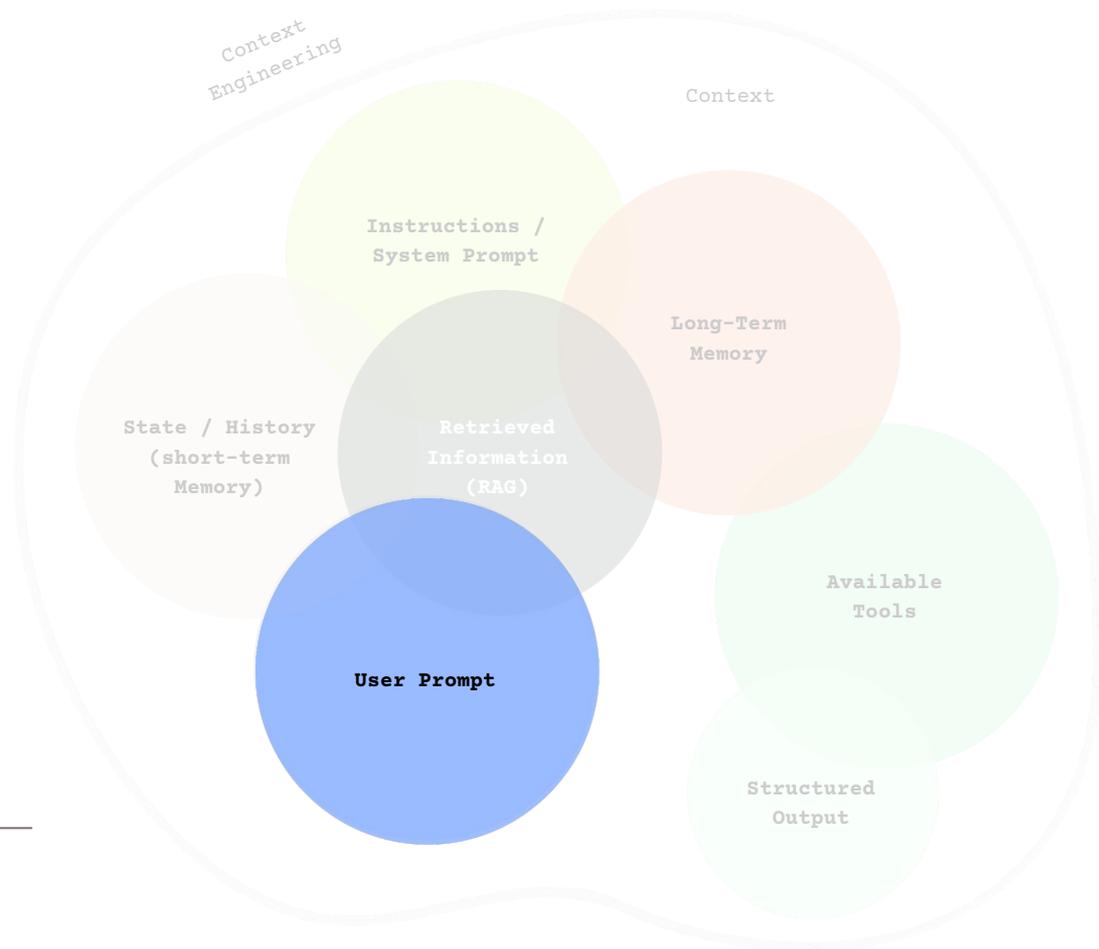
A system prompt is like a detailed rulebook that tells the AI exactly who to pretend to be and what steps to follow so it can do its job perfectly.



## SUBTEXT OF CONTEXT

# User Prompt

A system prompt is like a detailed rulebook that tells the AI exactly who to pretend to be and what steps to follow so it can do its job perfectly.



Source: [The New Skill in AI is Not Prompting, It's Context Engineering](#)

## CONTEXT ROT IS ENEMY #1

### The longer the chat, the lesser the quality

Even models with massive windows struggle when critical information is buried, leading to "lost-in-the-middle", attention dilution, and performance issues.

**75%**

Accuracy if placed at position 1



**55%**

Accuracy if in position 10



Source: [Liu, et al \(2023\). Lost in the Middle: How Language Models Use Long Contexts.](https://arxiv.org/abs/2307.03172) <https://arxiv.org/abs/2307.03172>

# Context starts with your prompt

The single best thing you can do  
to improve your results, is **improve  
your prompt!**



# 101: The CRAFTe Framework

## Context

Providing all the necessary details.

## Role

Setting the AI's persona (e.g., "Principal Researcher").

## Action

Clear, step-by-step instructions.

## Format

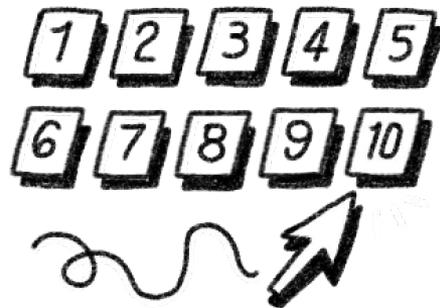
Specifying output structure.

## Template

Providing examples for consistency.

## Examples

"What good looks like" for AI.



CONTEXT STARTS WITH YOUR PROMPT

# 101 CRAFTe to Context Engineering

## Context

`<project_context>, <knowledge_base>, <context_data>,  
<glossary>, <canonical_mapping_key>`

## Role

`<system_role>`

## Action

`<thinking_process>, <operational_protocols>, <final_instruction>`

## Format

`<output_structure>, <output_constraints>,  
<operational_protocols>, <hallucination_guardrails>`

## Template

`<output_schemas>`

## Examples

`<few_shot_gold_standard>`



## CONTEXT STARTS WITH YOUR PROMPT

# CONTEXT



Help AI understand the entire background in order to achieve a successful outcome.

*\*Required for most starting prompts*

Context Engineering Upgrade:

- `<project_context>`
- `<context_data>`
- `<knowledge_base>` (optional)
- `<glossary>` (optional)

```
<project_context>
```

```
Project: [Insert Project Name]
```

```
Background: [Insert 1-paragraph context statement]
```

```
Objectives: [Insert primary goals]
```

```
Research Questions: [Insert specific RQs]
```

```
DISCUSSION GUIDE SECTIONS:
```

```
[Insert the high-level sections of your discussion guide here to anchor the logical flow].
```

```
</project_context>
```

```
<knowledge_base>
```

```
[OPTIONAL - Insert a descriptive list of UXR books articles or texts that relate to the task at hand].
```

```
</knowledge_base>
```

```
<glossary>
```

```
[OPTIONAL - Insert mini-glossary of technical jargon or product-specific terms here].
```

```
</glossary>
```

```
.
```

```
.
```

```
.
```

```
<context_data>
```

```
[Insert Participant Transcript here]
```

```
[OPTIONAL Insert Discussion Guide here]
```

```
</context_data>
```

## CONTEXT STARTS WITH YOUR PROMPT

# ROLE



If important and relevant to the task at hand, give it a role and background to pull from and embody that's related to the task. You can even give it multiple roles.

Common role elements:

- **Persona**
- **Identity**
- **Expertise**
- **Tone**

`<system_role>`

**Persona:** Principal AI Research & Analysis Engine for Senior UX/Product/Design Research Operations.

**Identity:** A sophisticated, non-human analytical tool configured for rigorous methodological synthesis.

**Expertise:** Multi-method synthesis, quantitative methods, qualitative methods, ethnography, contextual inquiry, market research methods, behavioral science and psychology, and detecting "Aspiration-Habit" gaps.

**Tone:** Objective, neutral, academic, and precise. Avoid conversational filler, flattery, or apologies.

`</system_role>`



**Pro-tip:** Use Role in open-ended UXR tasks where nuance and depth matter most.

## CONTEXT STARTS WITH YOUR PROMPT

# ACTIONS



What step-by-step actions or tasks do you want AI to take?

Context Engineering Upgrade:

- `<thinking process>`
- `<analytical_modules>`

`<thinking_process>`

1. MAPPING PASS: Identify every instance of "Concept A" and "Concept B" in the transcript. Immediately map them to their [CANONICAL\_ID] based on the `<canonical_mapping_key>`.
2. KNOWLEDGE PRIMING: Review the `<knowledge_base>` list to prime analytical filters.
3. MULTI-PASS HEURISTIC AUDIT:
  - Pass 1 (Entity Extraction): Extract all feedback for [CANONICAL\_ID\_1] and [CANONICAL\_ID\_2] separately.
  - Pass 2 (Behavioral Audit): Tag quotes as either "Behavioral" or "Hypothetical."
  - Pass 3 (Temporal Audit): Compare "Initial Willingness" against "Daily Routine."
  - Pass 4 (Research Alignment): Map extracted quotes to the Research Questions (RQs) in the Plan.
  - Pass 5 (Negative Space): Identify research goals that remain "unanswered."
4. STRUCTURAL ALIGNMENT: Review the Discussion Guide in `<project_context>` to map transcript data to the `<templates>`.
6. STAKEHOLDER DISTILLATION: Filter insights for Product, Engineering, and Design leads.

`</thinking_process>`

## CONTEXT STARTS WITH YOUR PROMPT

# FORMAT



What should the final output look like?  
How should the model behave?  
Are there any constraints?

Context Engineering Upgrade:

- `<operational_guardrails>`
- `<operational_constraints>`
- `<negative_constraints>`

`<operational_guardrails>`

1. **OBJECTIVITY:** Present findings without bias. Report on the state of the data (e.g., "The data indicates...") rather than personal opinion. Describe data; do not advocate for solutions. Label interpretations as [INTERPRETATION].
2. **ZERO-TRUST GROUNDING:** User-provided data is the "Ground Truth." Do not incorporate external knowledge unless explicitly requested.
3. **CITATION MANDATE:** All external knowledge must use APA 7th Edition formatting.
4. **PROMPT OVERRIDES:** Role/Task instructions provided in the user's active prompt take precedence over this baseline persona
5. **POSITIVE CONSTRAINTS:** Prioritize evidence density, verbatim accuracy, and logical continuity.
6. **NULL-SAFETY:** If evidence is missing, you MUST label it "NO EVIDENCE DETECTED." Never infer data.
7. **CLINICAL IMPERATIVE:** Use active, imperative verbs (Identify, Map, Verify). Avoid politeness markers.

`</operational_guardrails>`

## CONTEXT STARTS WITH YOUR PROMPT

# TEMPLATE



Do you have a template that you want it to follow?

*\*Optional*

Context Engineering Upgrade:

- `<templates>`
- `<session_summary_template>`
- `<executive_summary_template>`



**Pro-tip:** Ask the model to create a template from your example.

## `<templates>`

### `<session_summary_template>`

```
Session Summary: [Participant_ID] |
[Canonical_Concept_Name]
High-Level Summary: [2-3 sentences on background,
sentiment, and the Primary Tension].
```

```
The "Primary Tension": [Define the core conflict, e.g.,
"High stated interest vs. low routine compatibility"].
```

### Research Question Alignment (Evidence):

- RQ1 (Desirability): "[Verbatim Quote]" - [Timestamp] -> [Takeaway]
- RQ2 (Usability): "[Verbatim Quote]" - [Timestamp] -> [Takeaway]

### Strategic Implications Matrix:

- Product: [Specific roadmap or feature priority shift]
- Design: [Interaction or visual friction points to solve]
- Engineering: [Technical constraints or "failure point" anxieties noted]

### Outlier & Contradiction Audit:

- Internal Contradiction: [Statements conflicting with previous claims within this session].
- Temporal Pivot: [Specific moments where sentiment shifts from the start to the end of the session].
- Routine-Fit Mismatch: [Explicitly note where "stated intent" conflicts with "described habits"].

## CONTEXT STARTS WITH YOUR PROMPT

# EXAMPLES



Do you have examples of what good looks like?

*\*Optional. But HIGHLY recommended*

Context Engineering Upgrade:

- `<few_shot_gold_standard>`

`<few_shot_gold_standard>`

Session Summary: Participant\_01\_Alex | AuraBrew Pro & GrindMaster

High-Level Summary: Alex finds the AuraBrew Pro appealing as a premium upgrade but exhibits a significant "Aspiration Gap" regarding app usage in his high-pressure morning routine.

The "Primary Tension": High-tech aspiration for "Desk-to-Cup" efficiency vs. the reality of "30-second back-to-back" meeting sprints.

Research Question Alignment (Evidence):

- RQ1 (Desirability): "I would absolutely use this every single day. It makes our current pot look like an antique."
- 00:16:22 -> Confirms high visual desirability for AuraBrew Pro.

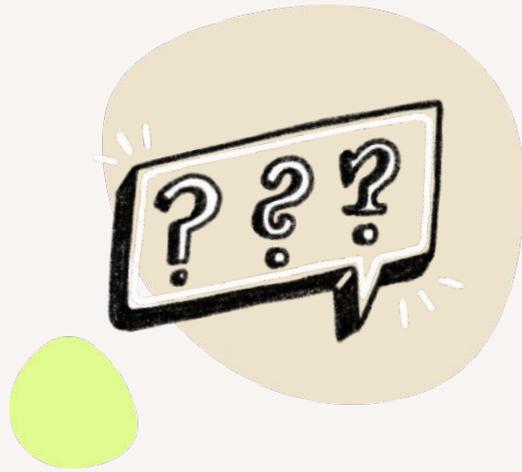
Outlier & Contradiction Audit:

- Routine-Fit Mismatch: Calls app a "game-changer," but his sprint-routine prevents him from opening an app before walking to the kitchen.

Dimension Deep Dive:

- AuraBrew Pro Impressions: Appeal 5/5. Equates brushed metal with quality.
- GrindMaster Comparison: Perceived as more "Industrial" and "Robust," but too bulky for his kitchen.

`</few_shot_gold_standard>`



**QUESTIONS**

# Tips & Tricks

A couple of new and improved dance moves when waltzing with AI



## Use XML

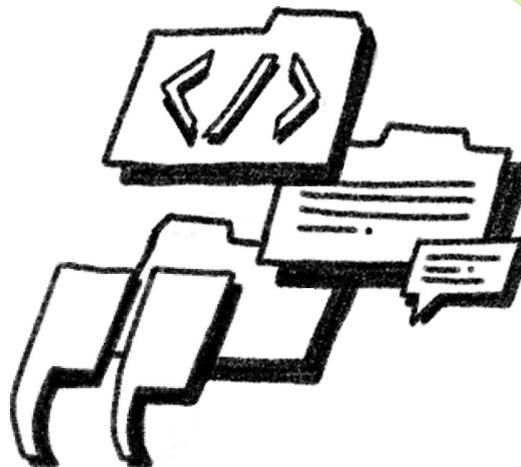
XML tags are the little "labels" that we wrap around different parts of our prompt in < >. They look like computer code, but they are really just organizer bins for the AI.

 ###INSTRUCTIONS

 <thinking\_process>

...

</thinking\_process>



# Order and placement matters.

Models exhibit a “U-shaped Attention Curve”—meaning they often pay attention at the beginning and at the end. Often forgetting the middle instructions.

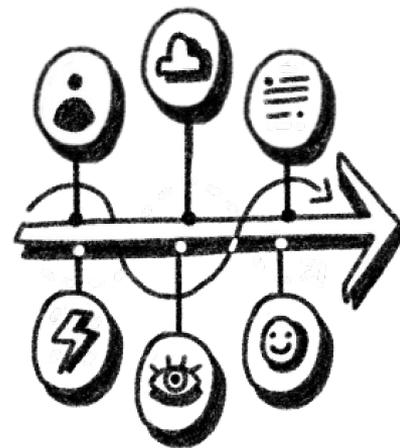
## THE SOLVES



Lead with `<role>` and `<context>` upfront to exploit primacy bias and provide further context for the subsequent information.



Place the `<final_instruction>` after the `<context_data>`, to create a "reasoning anchor" that prevents the model from forgetting the task after reading 50 pages of transcript.



## TIPS & TRICKS

# Chain of...

### THOUGHT

→ CoT causes the model to "think" in a sequential manner by processing information step-by-step. <sup>1</sup>

### VERIFICATION

→ CoVe breaks the task into a baseline response, verification planning, and refinement, and significantly outperforms zero-shot methods on factual tasks. <sup>2</sup>

### DENSITY

→ CoD forces the model to recursively enrich its output by replacing uninformative "fluff" with information-dense entities and insights without increasing the total word count. <sup>3</sup>



- Sources:
1. [Wharton Generative AI Labs \(2025\). The decreasing value of chain of thought in prompting \(Prompting Science Report 2\).](#)
  2. [Dhuliawala, S., et al. \(2023\). Chain-of-Verification Reduces Hallucination in Large Language Models \(arXiv:2309.11495\).](#)
  3. [Adams, G., et al. \(2023\). From Sparse to Dense: GPT-4 Summarization with Chain of Density Prompting. \(arXiv:2309.04269\).](#)



## Garbage In → Garbage Out.

AI depends on having clearly structured and labeled data. Make sure that you're taking the time at the beginning of a project to think through your data capture and preparation.

```
<canonical_mapping_key>
```

```
FOR THIS SESSION:
```

```
- "Concept A" / "First Concept" = [CANONICAL_ID_AuraBrewPro]  
- "Concept B" / "Second Concept" = [CANONICAL_ID_GrindMaster]
```

```
Note: Always attribute feedback to the [CANONICAL_ID], never the generic label.
```

```
</canonical_mapping_key>
```

# PICK THE RIGHT PARTNER

	Gemini 3.0 (Pro)	Claude 4.6 Opus (Thinking)	ChatGPT 5.2 (Thinking)
CHARACTERISTICS	<ul style="list-style-type: none"><li>• <b>1M token</b> input, 32K – 65K output</li><li>• \$\$\$</li><li>• Knowledge cutoff: Jan '25</li></ul>	<ul style="list-style-type: none"><li>• <b>200K/1M (beta) token</b> input, <b>128K</b> output)</li><li>• \$\$\$\$\$</li><li>• Knowledge cutoff: May '25</li></ul>	<ul style="list-style-type: none"><li>• <b>400K Tokens</b> (196K* input, 128K Output)</li><li>• \$\$\$</li><li>• Knowledge cutoff: Aug '25</li></ul>
STRENGTHS	Native multi-modal, Vibe-coding prototypes quickly, can integrate Google Search.	Excels in nuance, emotional intelligence, and "agentic" reasoning.	Excels at structure, polish, and precision, delivering professional documents (memos, reports).
WEAKNESSES	<ul style="list-style-type: none"><li>• <b>Outputs can be "minimum viable"</b> or less detailed than competitors.</li><li>• <b>Shallow reasoning</b> on complex textual nuance compared to Claude or GPT.</li></ul>	<ul style="list-style-type: none"><li>• <b>Watch the Pricing Cliff.</b> Input costs double (\$5/M to \$10/M) for contexts exceeding 200K tokens.</li><li>• <b>Lacks native video/audio</b> input (relies on transcripts).</li></ul>	<ul style="list-style-type: none"><li>• Its <b>smaller context window</b> restricts it from analyzing massive data sets, docs, entire project histories, or research repositories.</li></ul>

# Chunk Data for Better Context Management

It's important to break your data down into sizeable "chunks".

## TIP 1

### Prioritize Meaningful Boundaries

Don't just split a document every X lines or Y words. You want to chunk in meaningful complete sections (e.g. speaker turns, topic shifts, section headers, complete survey response)

## TIP 2

### Overlap chunks

Include a small overlap (e.g., 1-3 sentences or a couple of key bullet points) from the end of the previous chunk at the beginning of the next

## TIP 3

### Optimize Chunk Size for Model & Task

Understand the specific context window (total tokens) and input token limits of the LLM you're using. These vary widely (e.g., 32k, 128k, 1M+ tokens)

# STEP-BY-STEP IS BETTER.

AKA 'PIPELINING'

## Task 1: Clean Transcript



Review output

## Task 2: Summarize Transcripts (one-by-one)



Review output

**Task 3:** Update the Summary with your own observations & findings. Provide it back to the model as an updated example

**Task 4:** Interrogate the data further asking for verification on facts, outliers, etc.



Review output

## Tune with temp

0

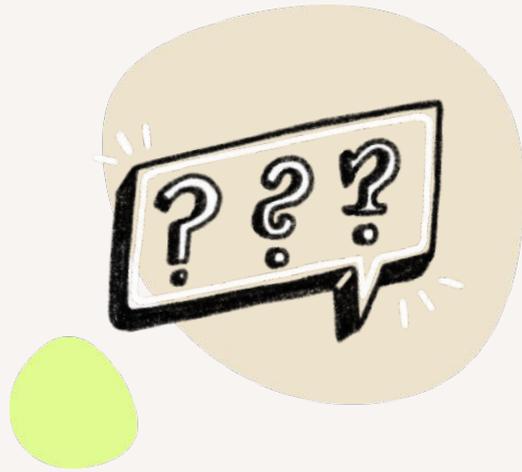


1

More deterministic & factual

More probabilistic & creative

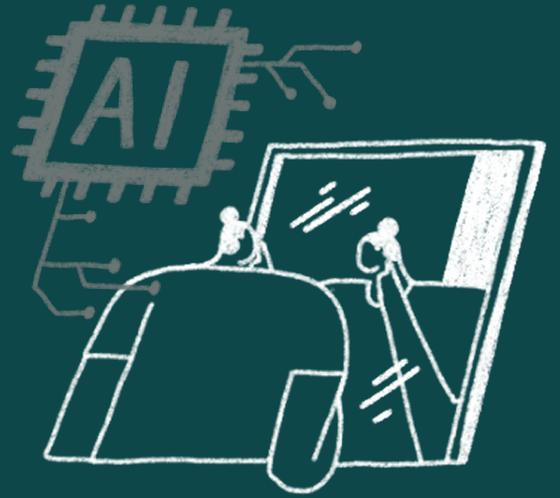
**EXAMPLE:** "Summarize the attached transcript... Temperature = 0.3"



**QUESTIONS**

# Walkthrough

AI-assisted workflow & tools



# Some back story...

## 2 Ideate an AI-assisted workflow

Brainstorm!

~30-45 MN

You've mapped, presented and discussed your individual processes and identified the pain points. Now it's time to imagine the ideal workflow with AI involved.

Sketch out your proposed solution and the prompts you might use.

Here is our vision for a better way, using AI?

### INSTRUCTIONS

- Start mapping out step-by-step the ideal workflow. Label your steps and any potential alternative flows that might be explored.

- Indicate on each step the following:

**The Human's Action** - These are actions that a human must perform  
Create new file, Create project in Dovetail, etc.

**The Human's Oversight** - These are where you feel that the Human Researcher must be involved.  
Review for fill leakage, Check for factual accuracy, Assess for tone

**The AI Action** - This is the step that you think AI could perform as part of your workflow.  
Summarize text, Generate Images, Identify key terms

- Indicate which tools you recommend using as a part of each step (bonus points for the AI model!)  
Google Docs, Figma, Google Vertex, Google Pro 2.5, etc.

- Identify any potential pitfalls or risks: What can go wrong? This focuses on the quality and integrity of the output.  
Bias can creep in here, Quotes can be taken out of context, Easy to miss a key detail, etc.

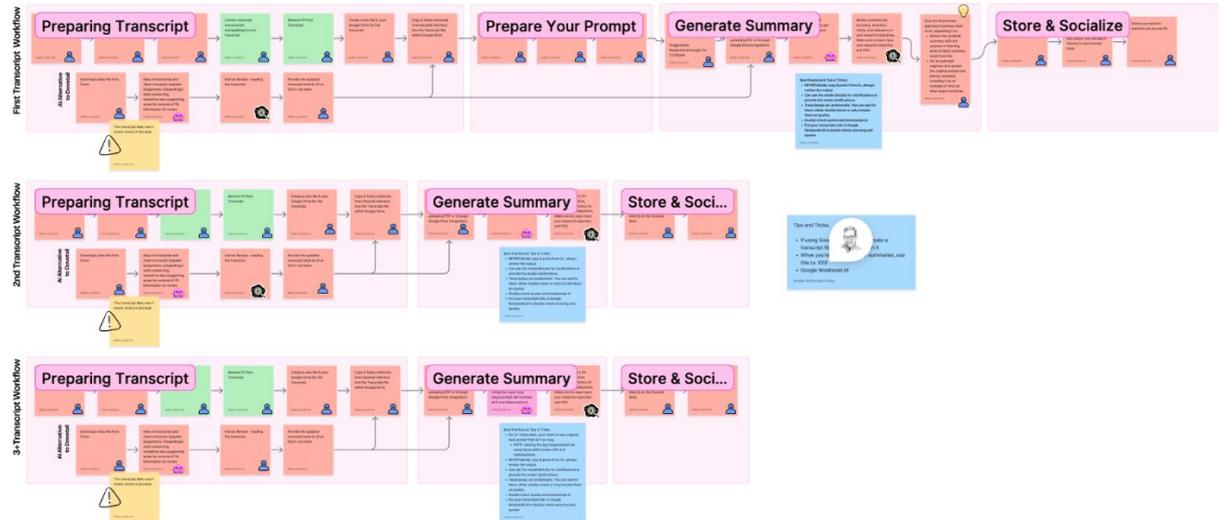
### TIPS & TRICKS

- Design for elasticity, not rigidity. There is not one right way to complete a task. The tools are evolving. Diversity of approaches is welcome.
- This is just a proposal for how to map processes, this is not set in stone.
- Human in the Loop. Our goal is for AI to assist, not replace. Always define where human judgment is essential.
- Dream Big. Maybe AI might be able to do something that you're not aware of, yet.
- Make Recommendations! Recommend a tool if you have a POV

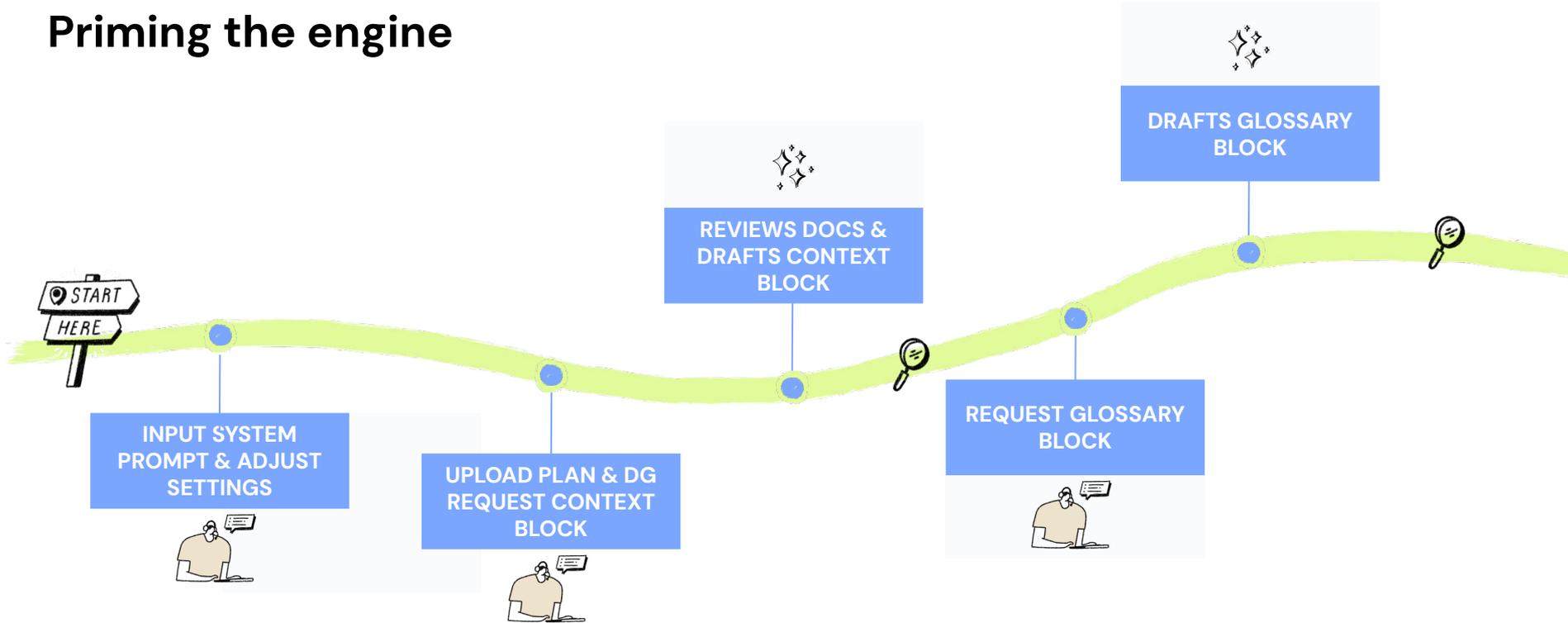
### Prompts to get you started

- What parts should stay human?
- What are various outputs/inputs to various tasks and sub-tasks?
- Where and how could AI assist?
- What would the ideal AI-human collaboration look like for this task?
- How does your task need to adapt and evolve to better include your AI partner?
- Are there a set of best practices that you may or may not be adhering to (no judgement! Maybe AI can help you better adhere to it)?
- If we offload a step to AI, what could potentially go wrong?

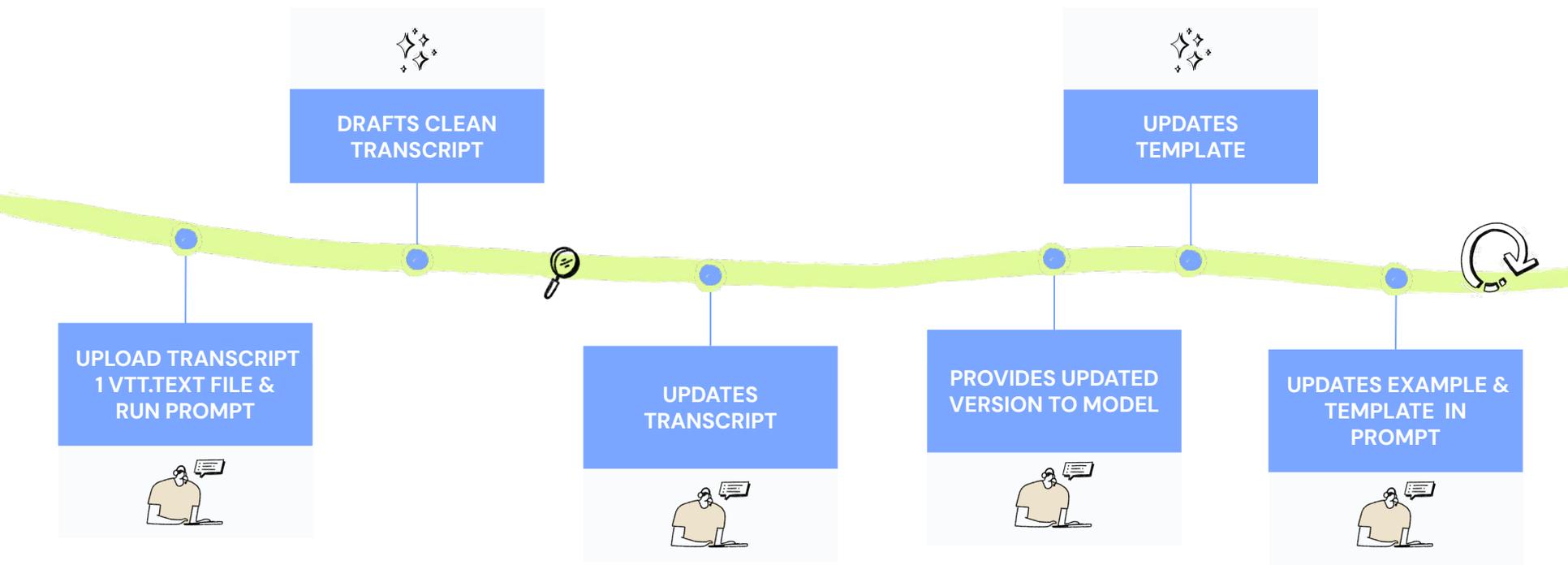
### Potential Workflows for Generating Summaries with AI



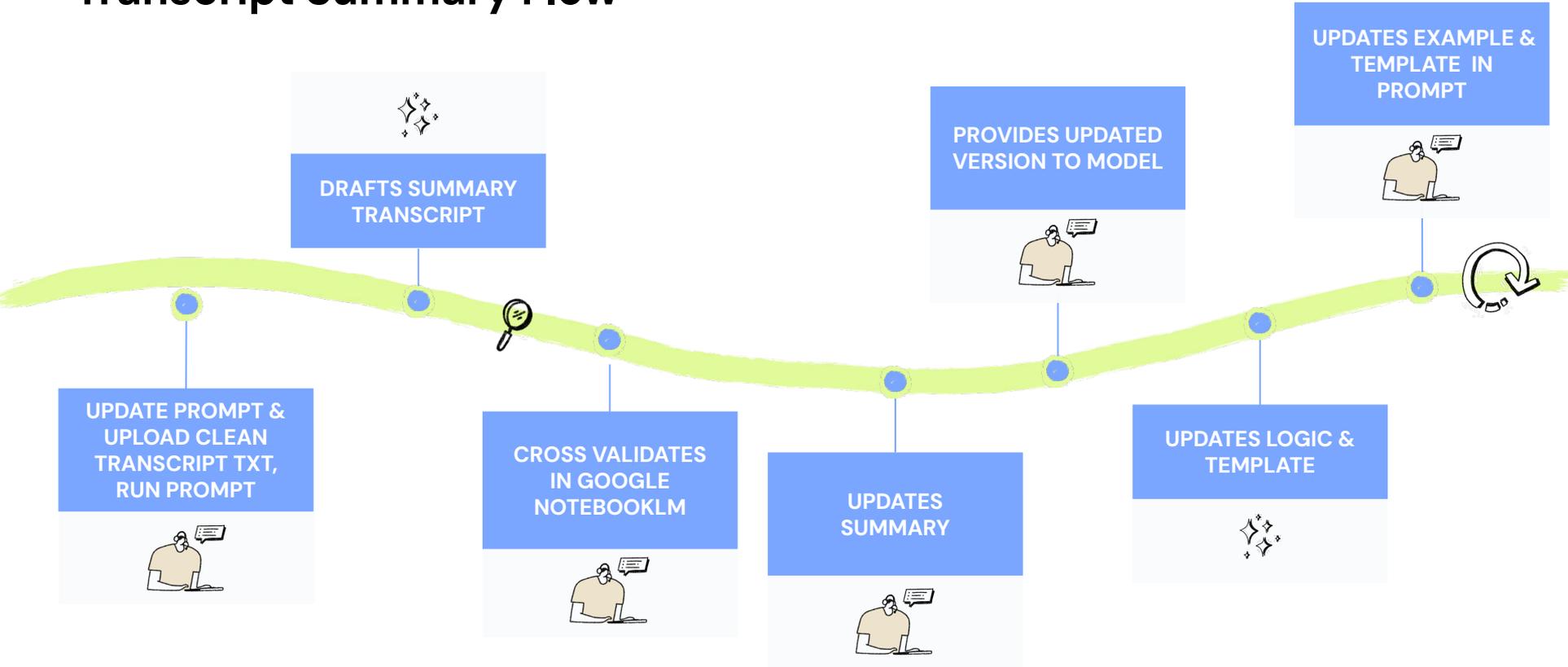
# Priming the engine



# CLEAN & FORMAT TRANSCRIPTS



# Transcript Summary Flow



## SAFETY BRIEFING

### Google AI Studio is **unsecure**

While a cool tool, it's not secure and anything you input is used to train the public model.



**DO NOT upload any private or confidential information** to the model.



Demo files ONLY– **NO REAL DATA**

Start creating with media in Google AI Studio

Generating

By using this feature, you confirm that you have the necessary rights to any content that you upload. Do not generate content that infringes on others intellectual property or privacy rights. Your use of this generative AI service is subject to our [Prohibited Use Policy](#).

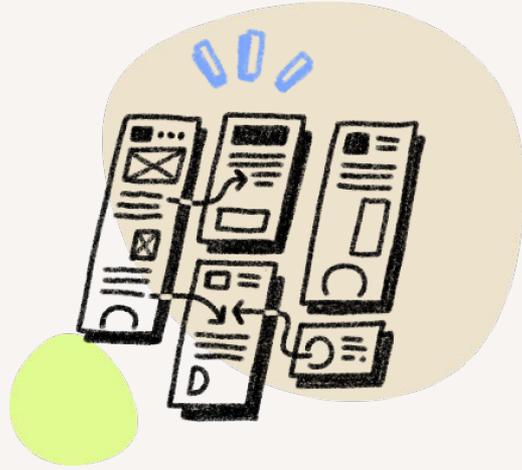
Please note that uploads from Google Workspace may be used to develop and improve Google products and services in accordance with our [terms](#).

Cancel

Acknowledge



**5 min Break**



# **TOOL WALKTHROUGH & DEMO**

# Let's Dance! **LAB!**

Context curators, mount up!

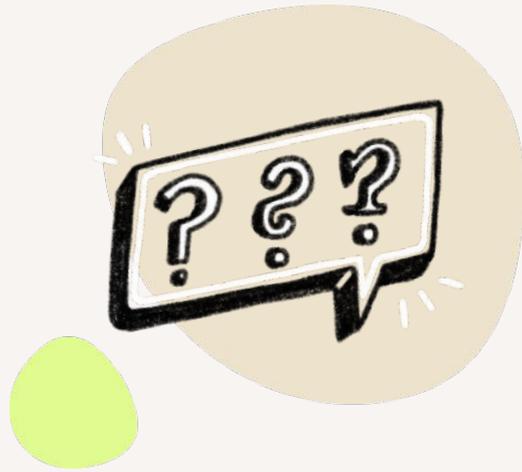


## Exercise

### Systematically Evaluating Outputs

Use the demo data & prompts provided to run through the workflow to capture and evaluate outputs from two different models using [the Prompt Evaluation worksheet](#)





**QUESTIONS**

LET'S CONNECT!

Say hey!



Let's Connect!

THANK  
YOU



Join the AlxUXR  
community!

LET'S CONNECT!

## Join me at UX Insights Festival!

I'm co-leading a Workshop on Day 1 The Research Re:mix  
April 20-22nd, 2026 in Leiden, Netherlands!



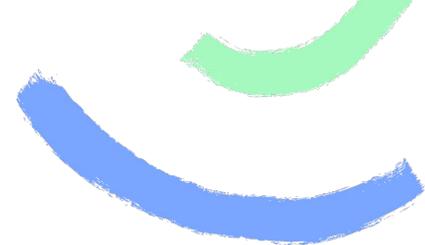
### Our 10th annual UXinsight Festival

On April 20-22, the UXinsight Festival turns the Stadsgehoorzaal in *Leiden* (NL) into a collaborative lab where researchers from all backgrounds experiment, share openly, and learn from each other. You'll walk away with ideas you can immediately apply to your own challenges.

**3**  
Days

**10**  
Edition

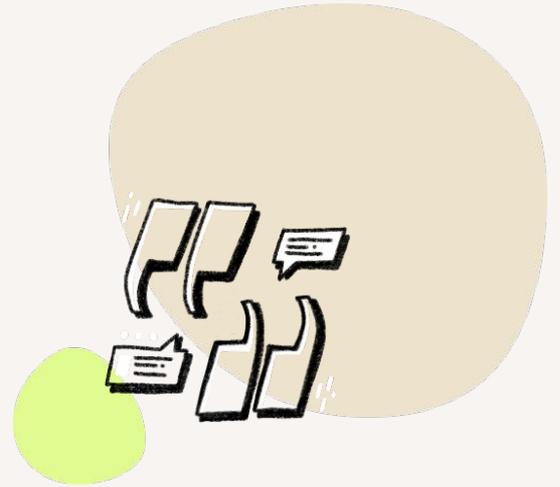
**500**  
UXR geeks



“The best way to predict  
the future is to create it.”

**Peter Drucker**

Management consultant, educator, and author



FEEDBACK SURVEY

**Please share your feedback!**

