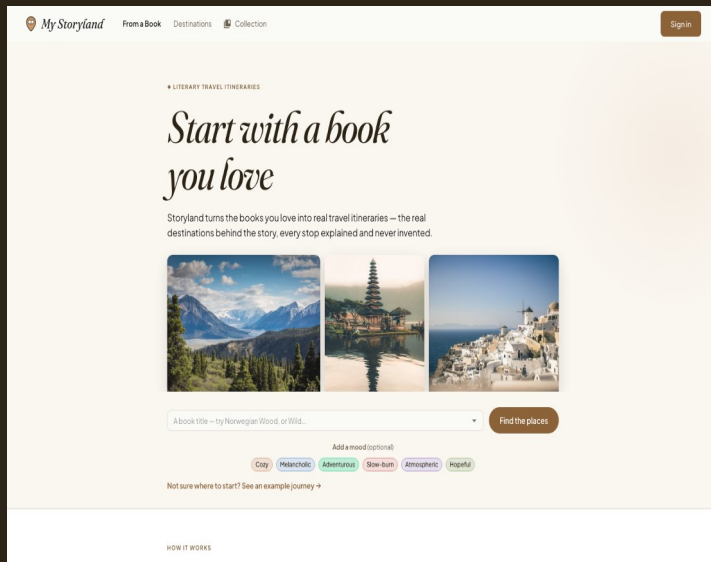
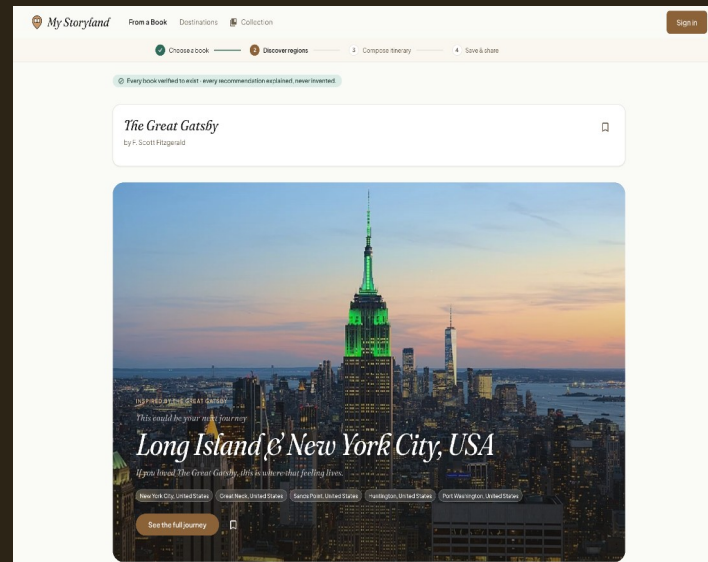
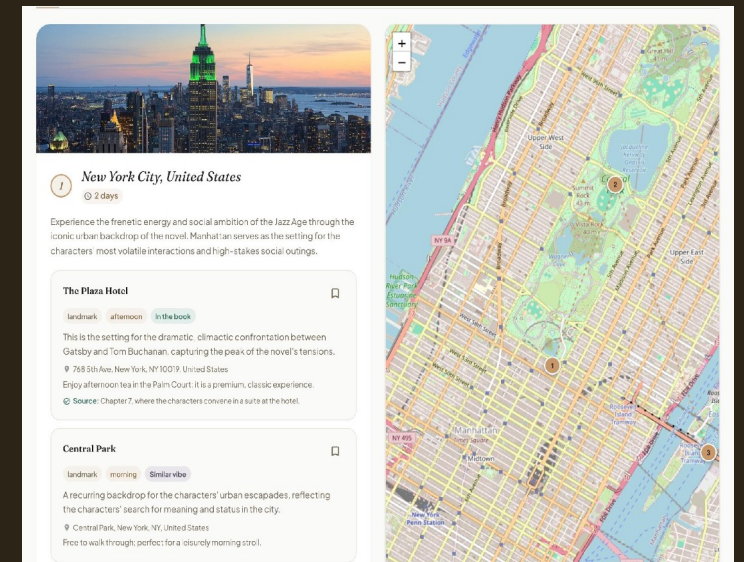




start with a book



pick a real region



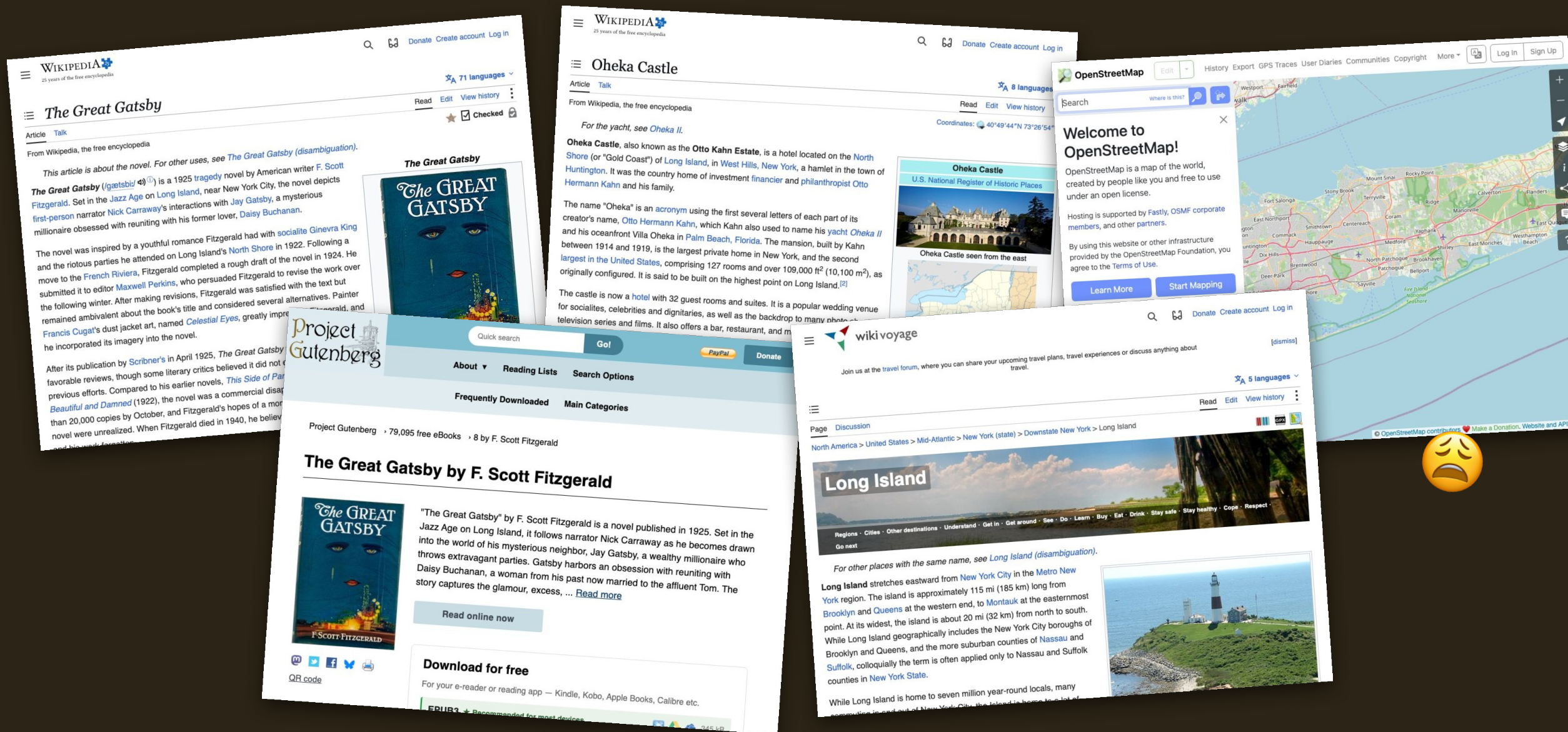
keep the itinerary

From Bookshelf to Borderline

Building a multi-agent AI that turns books into travel

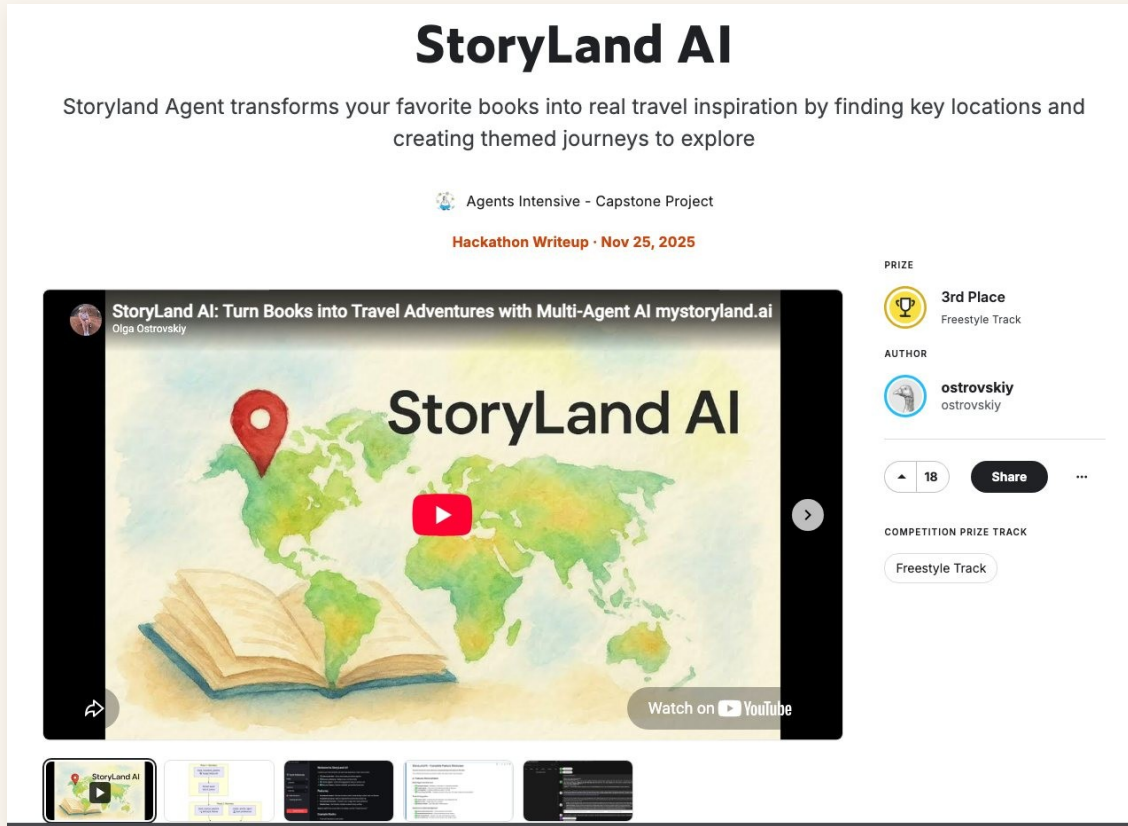
Olga Ostrovskiy · My Storyland · mystoryland.ai

DevFest DC 2026 · Fuse Center at George Mason · August 28



12 tabs. 0 itineraries.

From capstone to production



The writeup, still up on Kaggle

3rd place 5 days

Google AI Agents Intensive ·
Freestyle track

from idea to submission

- Nov 2025 Built during Thanksgiving week
- Jan 2026 3rd place, freestyle track
- Today live at mystoryland.ai

The user journey, in three steps

Atlanta

United States · 1 day

Margaret Mitchell wrote *Gone with the Wind* here and set much of its second half in a burning, rebuilding Atlanta. These are the real places behind the author and her story.

1 Margaret Mitchell House

Literary landmark Morning

The Midtown apartment Mitchell called "The Dump," where she wrote most of *Gone with the Wind* between 1926 and 1929. Now a museum to the novel and its author.

979 Crescent Ave NE, Atlanta, GA

2 Atlanta History Center

Museum Afternoon

Holds one of the largest *Gone with the Wind* collections anywhere — Mitchell's manuscripts, letters, and props from the 1939 film.

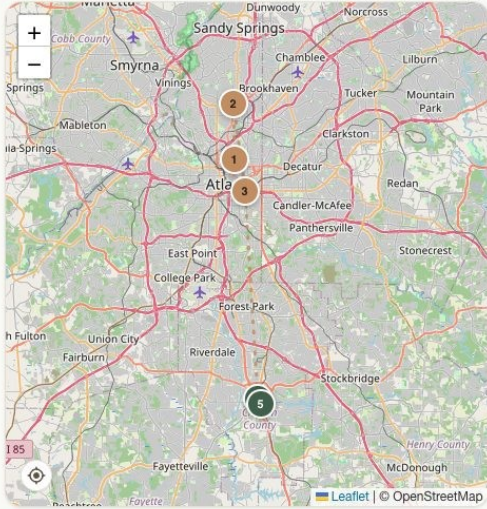
130 W Paces Ferry Rd NW, Atlanta, GA

3 Oakland Cemetery

Historic site Late afternoon

The Victorian garden cemetery where Margaret Mitchell is buried, and a vivid window onto the Civil War-era Atlanta the novel depicts.

248 Oakland Ave SE, Atlanta, GA



Jonesboro

United States · 1 day

Clayton County is the real-world countryside Mitchell drew on for Tara, the O'Hara plantation. The town has preserved the antebellum landscape the novel made famous.

4 Road to Tara Museum

Museum Morning

A museum devoted entirely to the novel and film, housed in Jonesboro's historic 1867 train depot in the county Mitchell used for Tara.

104 N Main St, Jonesboro, GA

5 Stately Oaks Plantation

Historic home Afternoon

An 1839 Greek Revival antebellum house — the kind of plantation home Mitchell pictured as Tara — open for tours of the Old South she wrote about.

100 Carriage Ln, Jonesboro, GA

- 1 Choose a book
- 2 Select a region
- 3 Get your itinerary

Live at mystoryland.ai

Why a prompt is not enough

HYPOTHETICAL FAILURE MODE

DAY 2 — WEST EGG

Tour **Gatsby's mansion** in **West Egg**.
After lunch, walk to the **Buchanan estate** in **East Egg** and end the day overlooking the bay.

WHY IT DOESN'T WORK

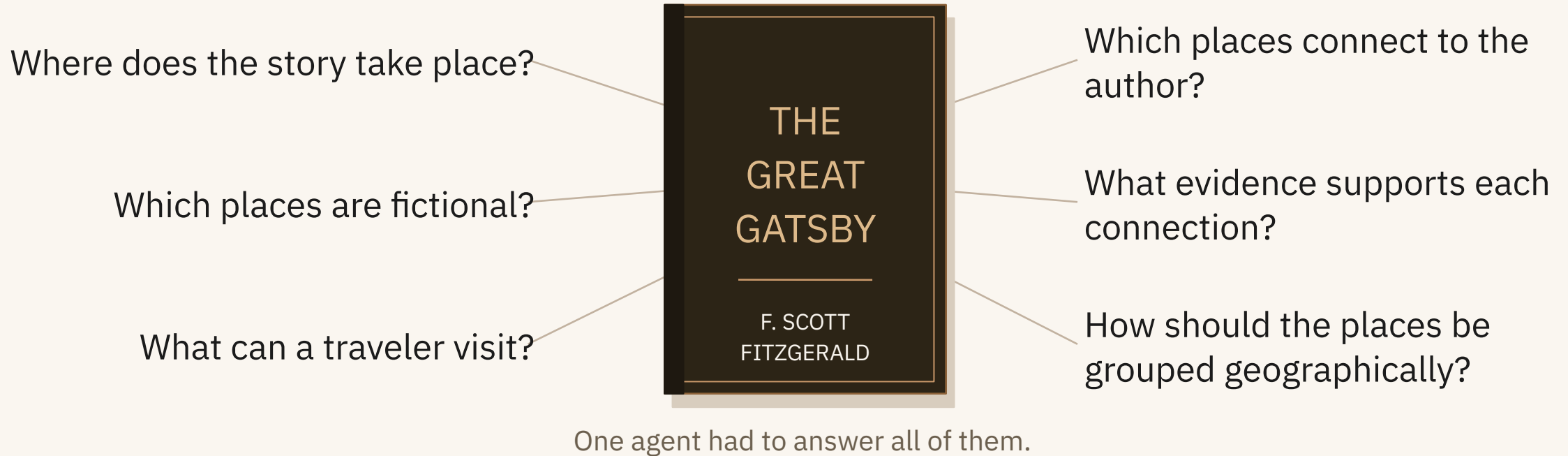
West Egg and East Egg are **fictional**.

There is **no verified** Gatsby mansion or Buchanan estate to visit.

The real Long Island connections require research and evidence.

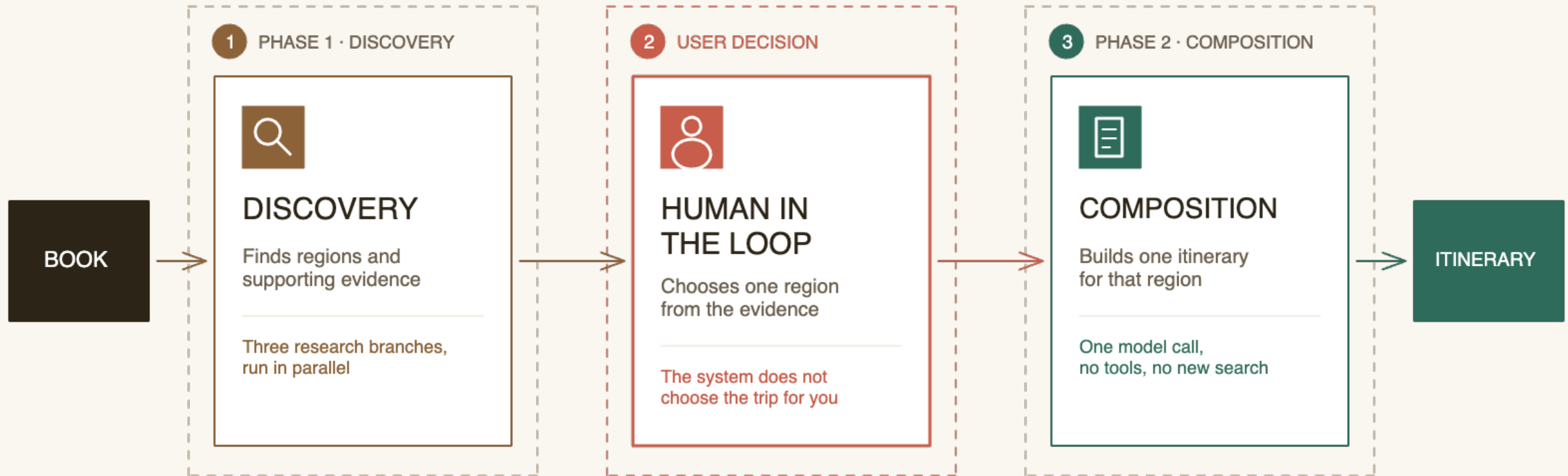
Connect fiction to real geography—without inventing the link.

The limits of one agent



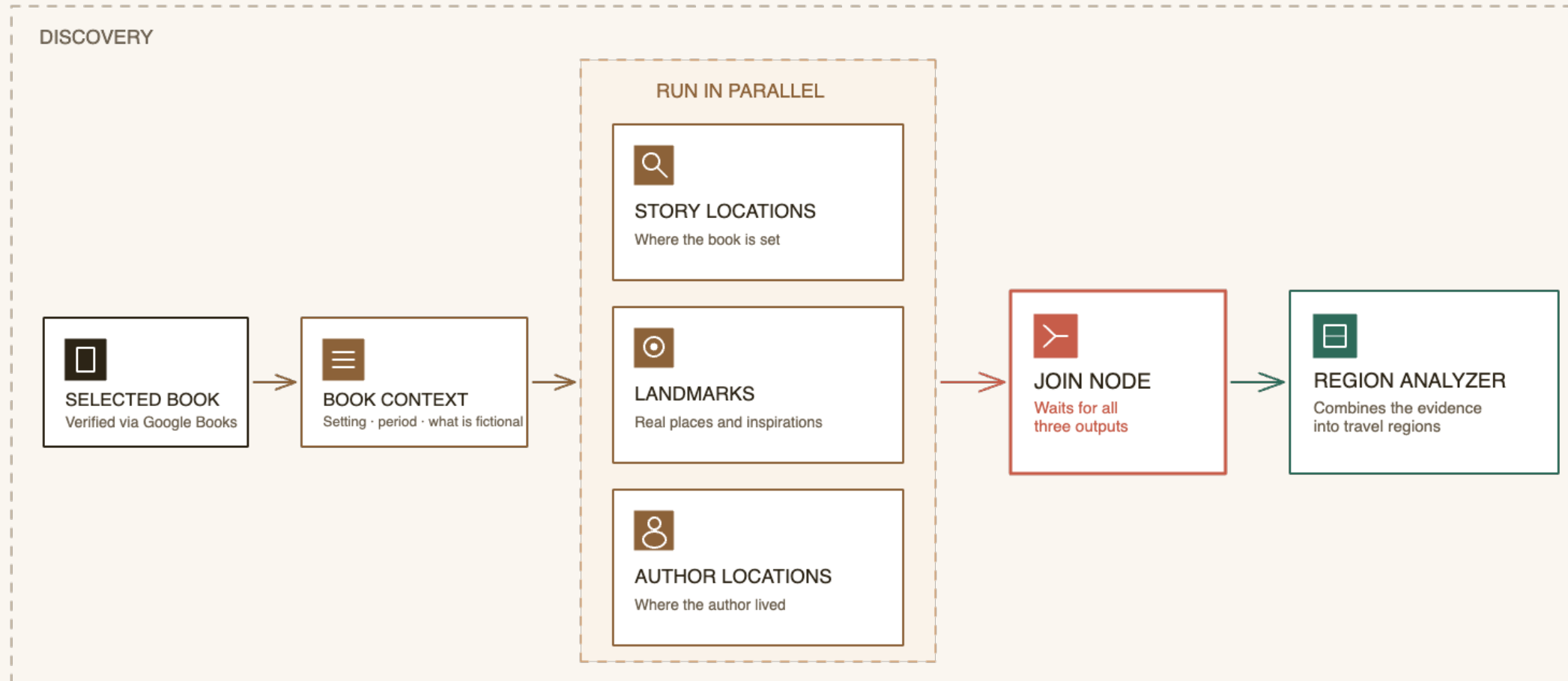
One agent could answer many—but not cover them all consistently.

The multi-agent workflow



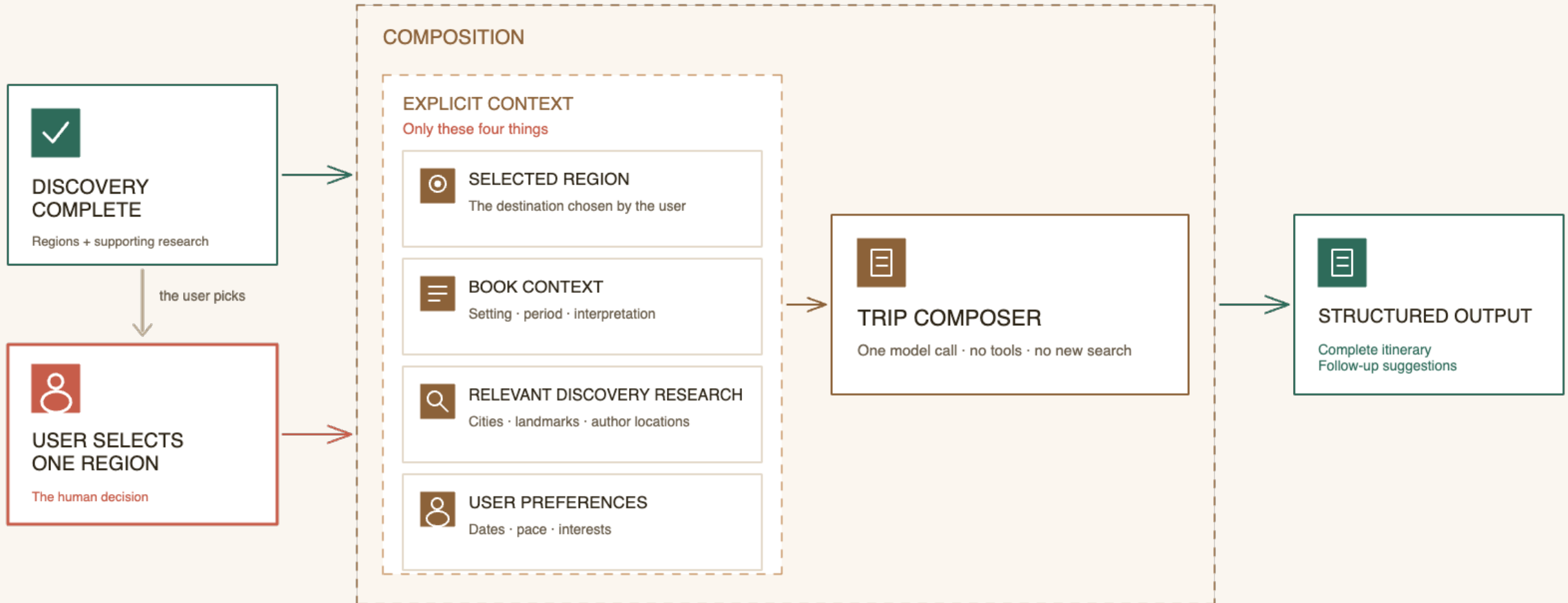
The system finds the options. The user chooses the trip.

Inside Discovery: parallel research



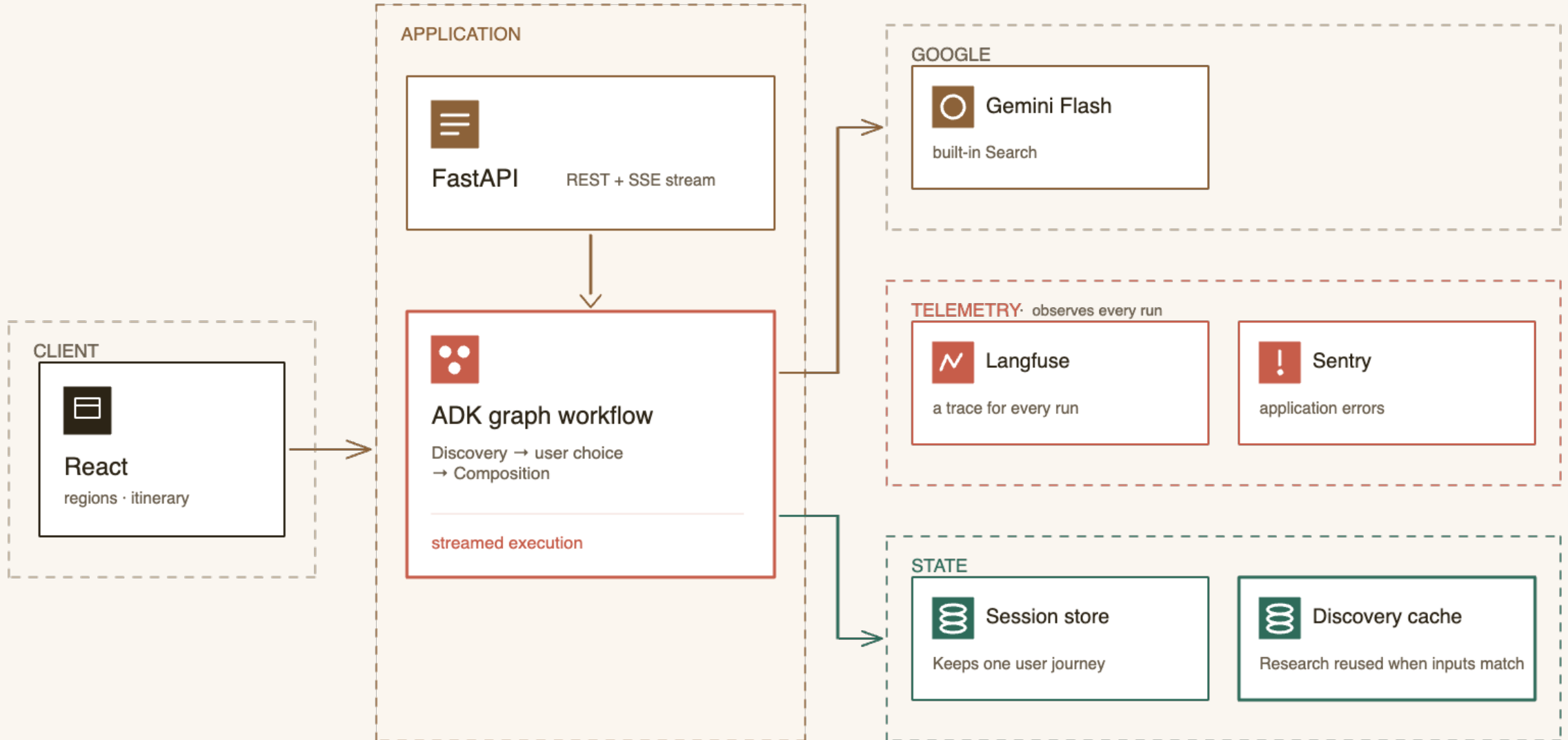
The analyzer starts only after all three research branches finish.

The handoff to Composition

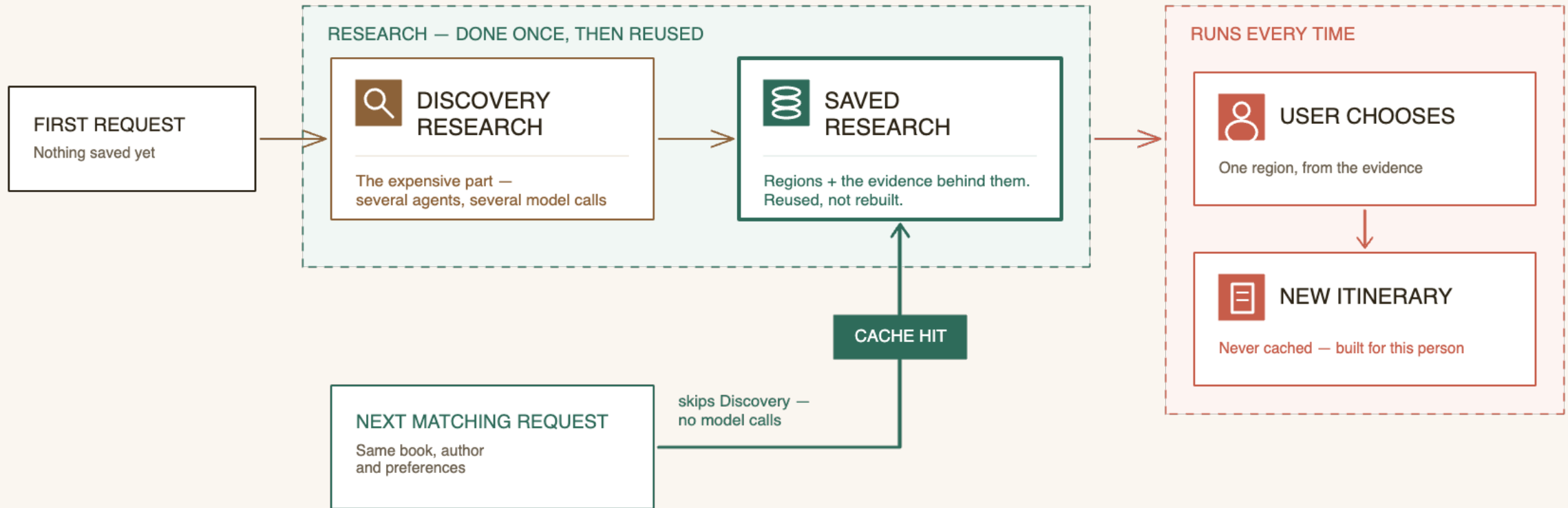


Explicit handoff: one region, confirmed research, preferences.

The application around the workflow



The Discovery cache



Reuse the research—not the final trip.

Deterministic checks in code

STRUCTURE

Is the result well-formed?

- Required fields are present
- Values have the right type
- Output has the expected structure

If not → reject before the next step

EXACT RULES

Does the result follow the rules?

- Strong claims have research
- Composition uses a region from Discovery
- The address matches the itinerary city

If not → downgrade, reject, or correct

Deterministic: same input, same decision, no additional model call.

Tracing a bad result

The screenshot displays the Langfuse interface for a production trace. The left sidebar shows a tree of trace nodes. The main panel shows the details of a selected node, 'gemini-3.1-flash-lite-preview_call', which is highlighted with a red box. The right panel shows the output of this call, which is a JSON object containing a text block and a list of author information.

```
output: {"model_version": "gemini-3.1-flash-lite", "content": {"parts": [{"text": "As an author biography specialist, I have compiled information regarding key sites connected to Fyodor Dostoevsky, the author of Crime and Punishment."}], "structured": [{"text": "1. Author's Hometown", "structured": [{"name": "Moscow (Mariinsky Hospital for the Poor)", "type": "Birthplace", "city": "Moscow, Russia"}]}, {"text": "2. Author Museums", "structured": [{"name": "The Dostoevsky Museum (Literary-Memorial Museum of F. M. Dostoevsky)"}]}}
```

This branch returned a result without Search evidence

Langfuse — one real production trace

One trace turns a bad answer into a debuggable path.

WHERE?

Which component introduced the claim?

WITH WHAT?

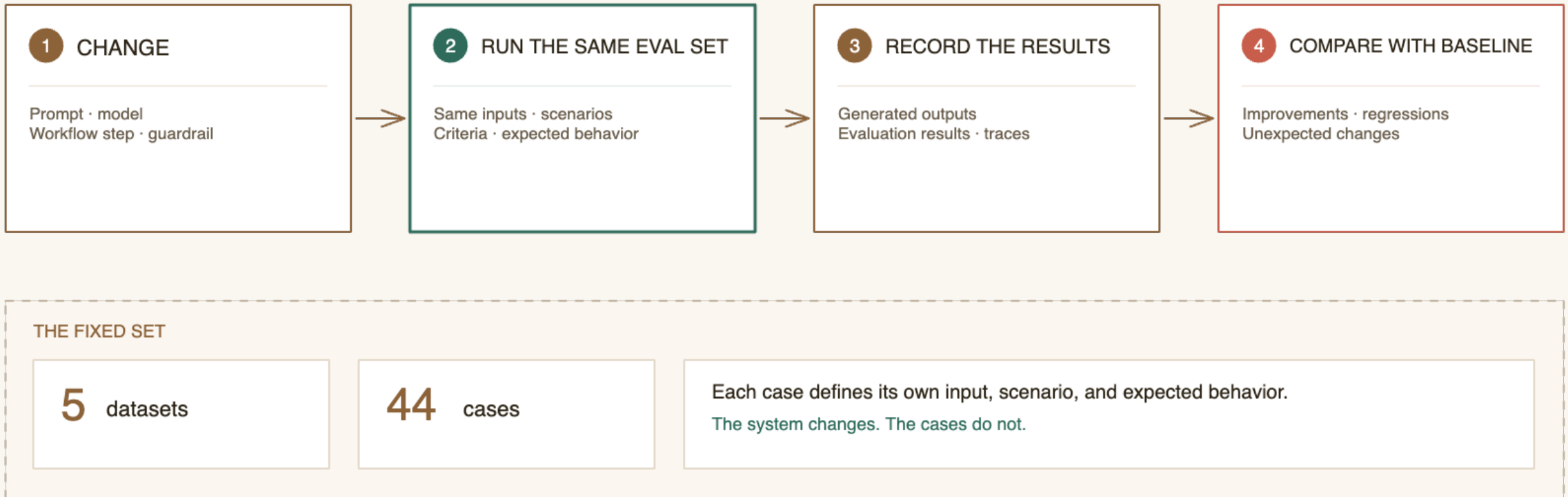
What context and evidence did it receive?

WHEN?

What happened before and after it?

Repeatable evaluation

The loop around every meaningful change.



The cases stay fixed, so changes become comparable.

Calibrating the LLM judge

LLM JUDGE = a separate model call that scores an output against a written rubric

ITINERARY CLAIM

“Portland Observatory – Oregon”

5/5 GEOGRAPHICAL_ACCURACY
5 = All locations accurate and real
the judge's own rubric · `evaluation/tools/llm_scorer.py`

FACT CHECK

Wrong state — it is in Maine

CALIBRATING THE JUDGE

After a judge-scored run, review a small random sample

- 1 Random cases · not only the low scores
- 2 Human review in the Langfuse annotation queue
- 3 Compare the human score with the judge
- 4 Re-anchor the rubric where they disagree

A judge score is a useful signal—not ground truth.

Human review calibrates the evaluator—not every production result.

Complete. Supported. Measured.

IS IT COMPLETE?

Wait for every required branch.

IS IT SUPPORTED?

Keep evidence with every strong claim.

CAN WE MEASURE IT?

Run the same cases—and check the judge.

A great book doesn't just end—it opens a door.

Helping a reader walk through that door became a systems problem.



TRY STORYLAND
mystoryland.ai



EXPLORE THE CODE
github.com/o-ostrovskiy/storyland-ai

Olga Ostrovskiy
DevFest DC 2026