

Agentic AI & Generative Tools for **Research**

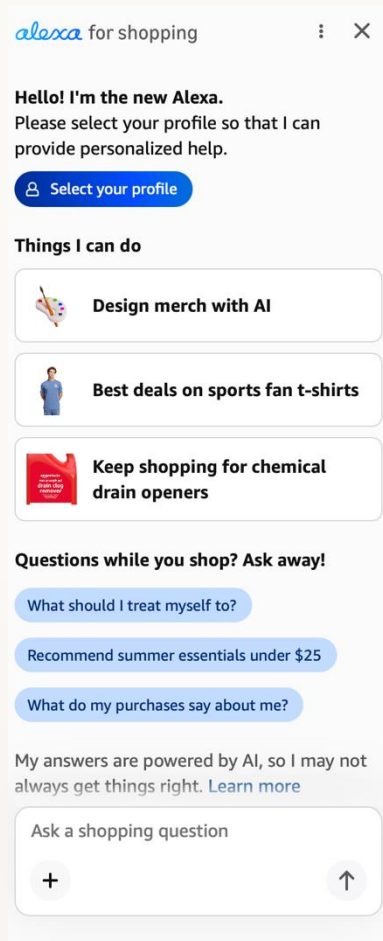
Designing Hypotheses with Deep Research, Troubleshooting Coding Local Pipelines (R & Python) and Logging Work

The Modern AI Toolkit

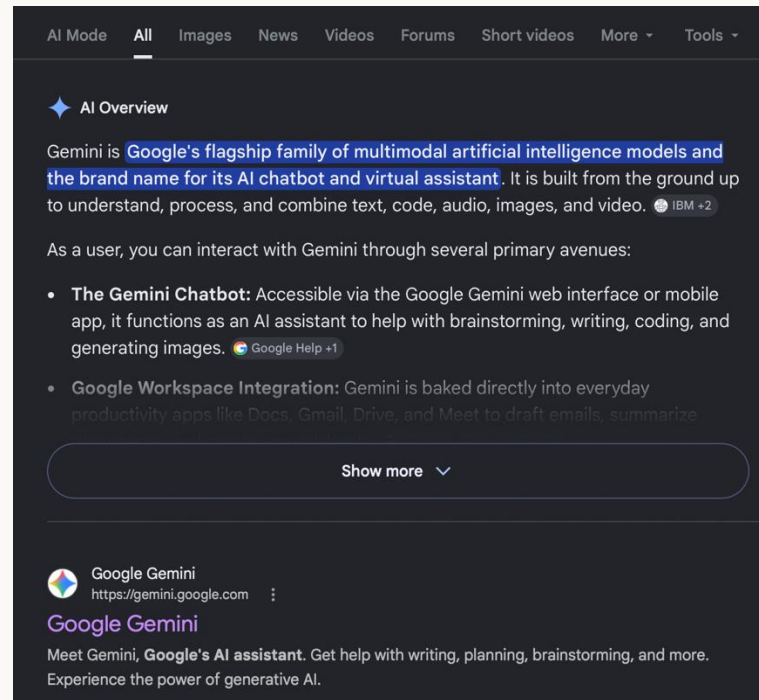
A transition from passive search bar queries to integrated, terminal-enabled agentic systems designed to build, analyze and troubleshoot scientific data

AI support is everywhere

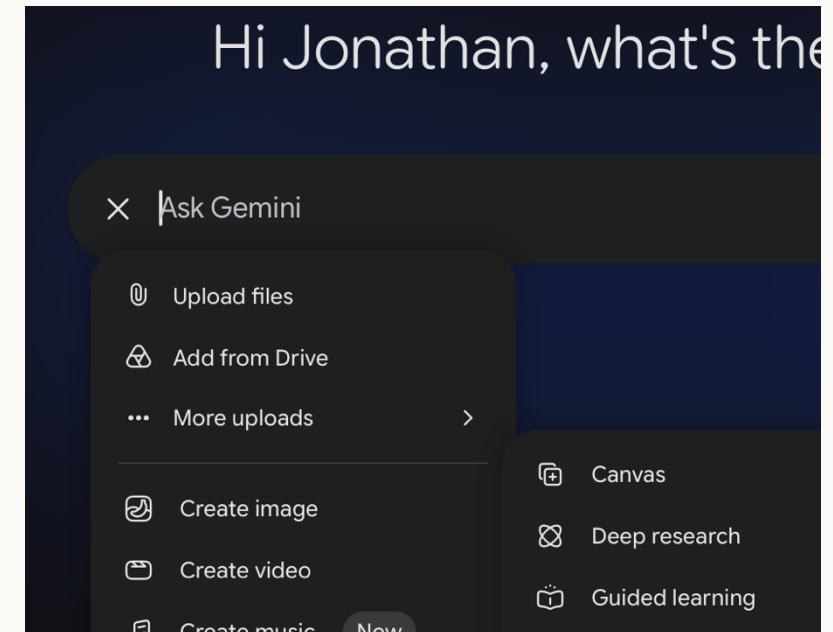
AI agents and regular life



Standard Google search



Beyond search to deep research



Two Modes of Gemini Discovery

Standard Gemini Chat

Fast, Single-Turn Synthesizer

- **Typical Latency:** Immediate response (< 5 seconds).
- **Core Use Cases:** Syntax lookup, code formatting, debugging single lines, fast biochemical definitions.
- **Generic Example:** *"Explain the molecular mechanism of CRISPR-Cas9 base editors."*
- **Case Study Option:** *"Explain what ice nucleation and the *inaZ* gene are."*

Gemini Deep Research

Autonomous Literature Mapping

- **Typical Latency:** Multi-minute, sequential deep queries (2-3 mins).
- **Core Use Cases:** Comprehensive literature surveys, synthesis of divergent preprints, assessing data trends.
- **Generic Example:** *"Analyze recent papers on adaptive metabolic traits in agricultural microbiomes."*
- **Case Study Option:** *"How does recombination rate vary across *Xanthomonas* lineages?"*

Claude Code vs. Google Gemini

Google Gemini Web Interface

The Analytical Research Companion

- **Operational Frame:** Web browser workspace environment.
- **Ecosystem Strengths:** Ideal for exploratory reading, generating conceptual outlines, and reviewing raw text blocks.
- **Primary Bottleneck:** Disconnected from files. Requires tedious manual copying, pasting, and folder configuration.

Claude Code (Terminal Agent)

The Command-Line Executor

- **Operational Frame:** Local terminal system shell interface.
- **Ecosystem Strengths:** Possesses direct file-system permissions, letting it write, read, edit, and troubleshoot scripts on the fly.
- **Primary Workflow:** Scaffolds code structures, initiates repos, and runs commands directly inside terminal workspace folders.

Actionable Research

Transform broad landscape literature queries into reproducible, local code executions:

- **Step 1: Research:** Leverage deep research models to synthesize the state of the field for a specific system.
- **Step 2: Create Concrete Objectives:** Prompt the model to extract exactly 3 testable scientific hypotheses from that literature overview.
- **Step 3: Ask Gemini for a Summary:** Ask the model to experiments
- **Step 4: Compare Deep Research to Regular Chat:** Prompt both Deep Research and also Canvas mode (with uploaded summary)
- **Step 5: Discuss with neighbor results**



Hands-On Practical Lab

Using Gemini to advance your research quickly, locally

Research visualization roundup

Instructions: Before writing scripts, ask Gemini to explain the core biology so you can build validation standards based on research.

Bacteria can make ice! We though found that some bacteria are losing this biophysical ability to nucleate water into ice. You need to visualize this. You are given a tree file that describes the phylogeny of a bacterium called *Xanthomonas translucens*. We also have metadata that can be used to annotate the figure with stain names, sequencing type, and *inaZ*_status. *inaZ* status is listed because of this gene's presence or absence (0=absent, 1=present).

You will first ask Gemini to explain the biology behind *InaZ* and also the bacterium *Xanthomonas translucens*. Then read this one-page summary and use it to help create and annotate your tree.

Why do you personally think that *InaZ* might be missing from some of these bacteria? What impact would that have on the biology based on your summary/deep research?

You can use whatever tool is best for you, but I recommend using R studio for fun.

Lab Files	Data Format	Description / Variables Tracked
xtt_rerun_core_tree.treefile	Newick (Tree)	Phylogenetic tree containing branch lengths and tip labels.
final_summary_with_all_metadata-3.csv	CSV (Metadata)	Tracks: sample ID, geographic location, date, and `inaZ` status.

Research visualization roundup

Instructions: Before writing scripts, ask Gemini to explain the core biology concepts, then build validation standards based on research.

Google Gemini

You are given a tree file that describes the
Hi Jonathan, what's the tree

Ask Gemini

Upload files

Add from Drive

More uploads

Create image

Create video

Create music

New

Canvas

Deep research

Guided learning

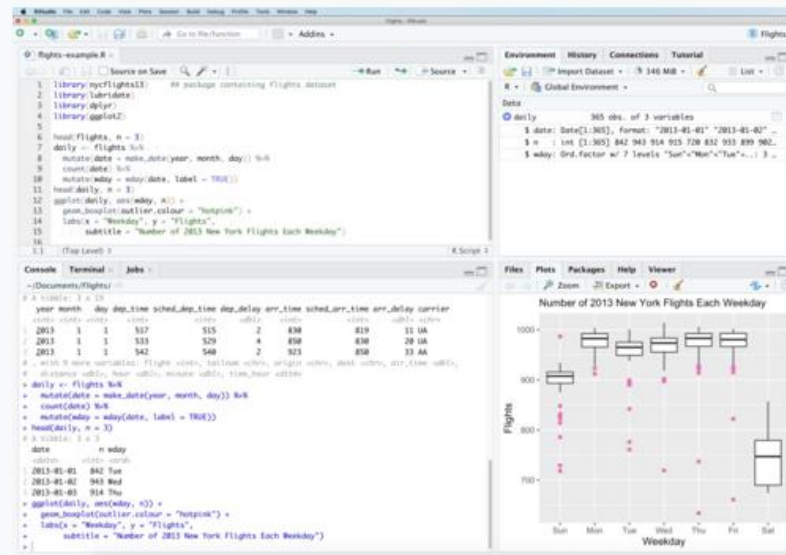
xtt_rerun_core_tree.treefile

final_summary_with_all_metadata-3.csv

CSV (Metadata)

Tracks: sample ID, geographic location, date, and `inaZ` status.

RStudio



slucens. We also have
_status. *inaZ* status is

as *translucens*. Then read

pect would that have on

ed

branch lengths and tip labels.

Research visualization roundup

Google Gemini

You are given a tree file that describes the
Hi Jonathan, what's the

Ask Gemini

Upload files

Add from Drive

More uploads

Create image

Create video

Create music

New

Canvas

Deep research

Guided learning

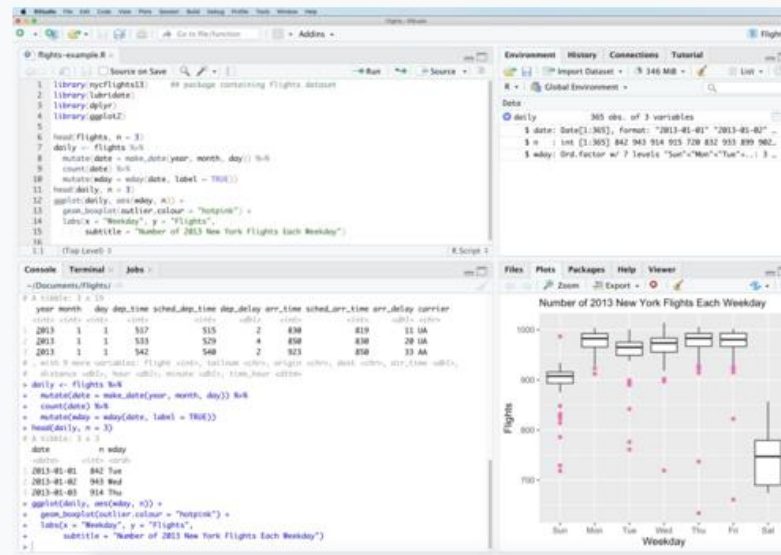
xtt_rerun_core_tree.treefile

final_summary_with_all_metadata-3.csv

CSV (Metadata)

Tracks: sample ID, geographic location, date

RStudio



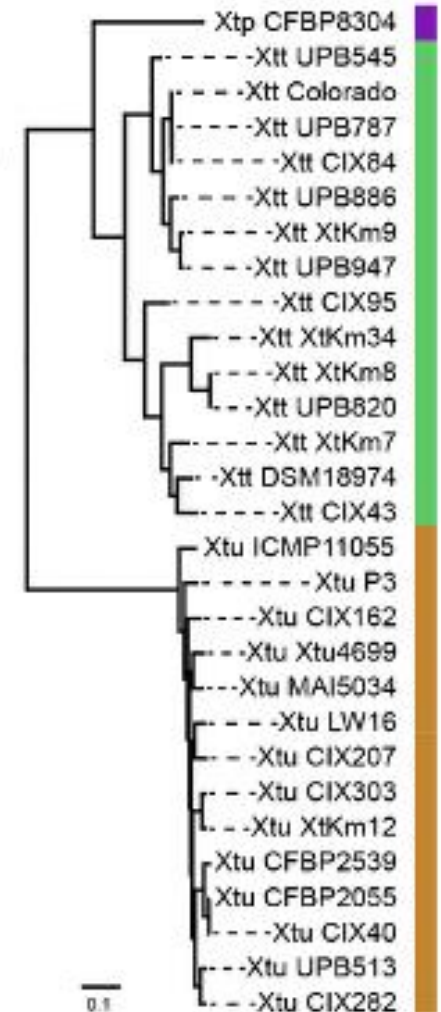
slucens. We
_status. ina

as transluce

pact would

ed

branch lengths



Exercise: Interactive Prompt Refinement

Broad Prompt (Low Actionability)

```
"Help me write some code to analyze my tree file and merge it with my data sheet."
```

- **Why it fails:** Model doesn't know file names, column labels, coding language, or library preferences.
- **Result:** Generic code that throws path and column mismatch errors.

Refined Prompt (High Actionability)

```
"Write an R script using 'tidyverse' and 'ape' to join 'my_tree.treefile' tip labels with 'sample_id' in 'data.csv'."
```

- **Why it succeeds:** Declares specific language, target libraries, file paths, and exact merge column labels.
- **Result:** Ready-to-run code tailored directly to your file structures.

Exercise: Interactive Prompt Refinement

Broad Prompt (Low Actionability)

```
"Help me write some code to analyze my tree file and merge it with my data sheet."
```

- **Why it fails:** Model doesn't know file names, column labels, coding language, or library preferences.
- **Result:** Generic code that throws path and column mismatch errors.

Refined Prompt (High Actionability)

```
"Write an R script using 'tidyverse' and 'ape' to join 'my_tree.treefile' tip labels with 'sample_id' in 'data.csv'."
```

- **Why it succeeds:** Declares specific language, target libraries, file paths, and exact merge column labels.
- **Result:** Ready-to-run code tailored directly to your file structures.

Discover Prompt (Practical Actionability)

Why you might need this: Use your prompt to tell a narrative just like you would tell a person. Try to use this as an opportunity to mentor the agent.

I am a scientist working on how plant pathogens evolve and the mechanisms that lead to microbe emergence. I have created a tree file and a metadata file but I need help using R studio to make a tree that is annotated. Can you help me with this?

Questions & Discussion