INESC TEC

INSTITUTO DE ENGENHARIA
DE SISTEMAS E COMPUTADORES,
TECNOLOGIA E CIÊNCIA


# GenAI Workshop Series: Research Design & Execution

Yassine Baghoussi, Inês Sousa, Lia Patricio and Gabriel David

Auditorium A - INESC TEC

# GenAI Workshop Series - Session 2

## Designing with Intelligence

## From scattered ideas to rigorous, organized research

# Using AI as Your Research Assistant

Session 1 – Ideation & Planning: From Ideas to Ethics

- November 6, 2025 | 9:30 AM – 11 AM | Auditorium A

**Session 2 – Research Design & Execution: Designing with Intelligence**

- **January 16, 2026 | 9:30 AM – 11 AM | Auditorium A**

Session 3 – Analysis & Interpretation: Turning Data into Insight

- February 12, 2026 | 9:30 AM – 11 AM | Auditorium A

Session 4 – Reviewing & Dissemination: Sharing with Impact

- March TBD

# Today's Journey

90-Minute Workshop Structure:

| Time | Section | Focus |
|---|---|---|
| 10 min | **Part 1:** Foundation | In-Context Learning & Custom GPT |
| 25 min | **Part 2:** AI Scientist | Start Ready from Scratch |
| 15 min | **Part 3:** AI Task Management | Persistent Contexts, Documentation, Organization |
| 25 min | **Part 4:** Hands-On Exercise | Group AI in Research critique |
| 15 min | **Part 5:** Discussion & Wrap-up | Sharing insights, key takeaways |

# Session 1 Recap

What We Covered Last Time:

- **The AI Revolution in Research Discovery**
  - From keyword search to semantic understanding
  - Tools that find papers by meaning, not just matching terms
- **Key Principles from Session 1:**
  - AI as brainstorming partner, not oracle
  - Always verify AI-suggested citations
  - Use specialized tools for specialized tasks (Elicit, Scite.ai, connected papers etc..)
  - Combine tools for comprehensive literature coverage
- **The Gap We Left:** *You have ideas. You have literature. Now what?*

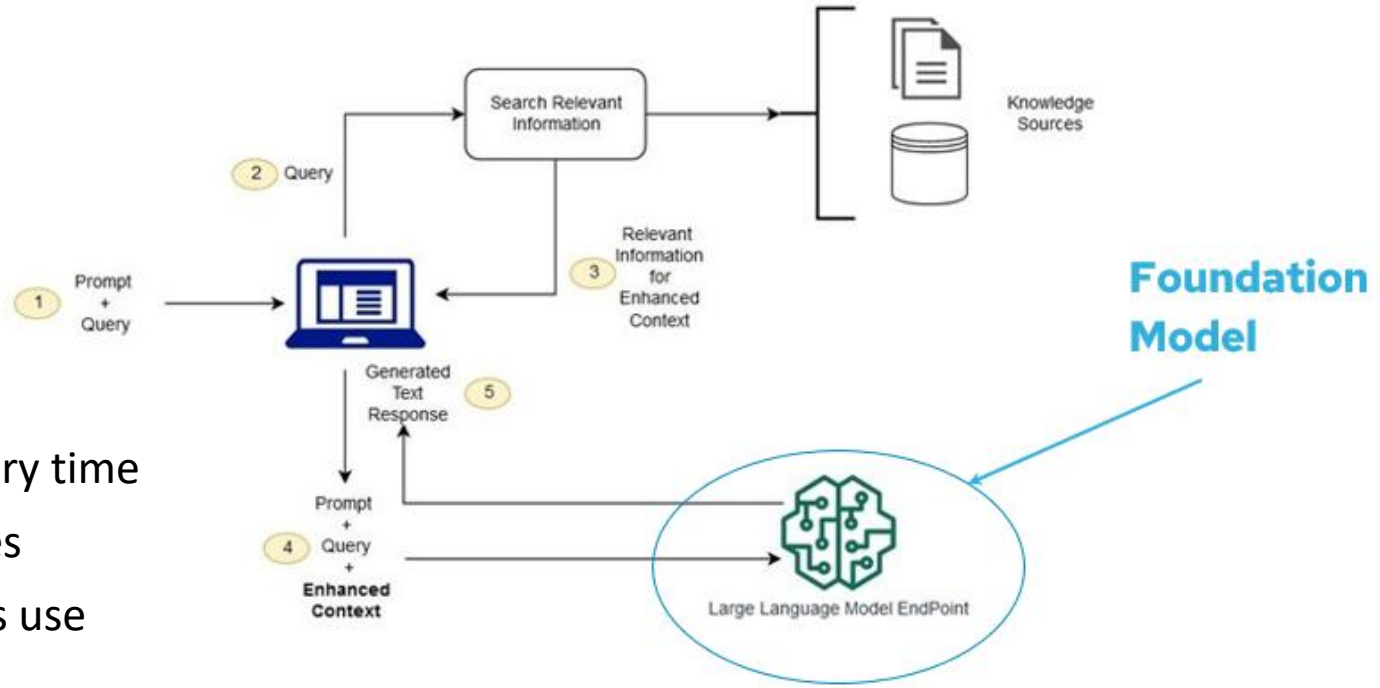
# Opening Quote

"The difference between these futures isn't in the AI, it's in how we choose to use it... we should aim to make these tools make us more **capable**, not just more **productive**.“

# PART 1: FOUNDATION

# Quick Recap: How LLMs Access Knowledge

We covered in Session 1:




RAG (Retrieval-Augmented Generation)

- Retrieves external documents at query time
- Grounds responses in specific sources
- What tools like Perplexity, Consensus use

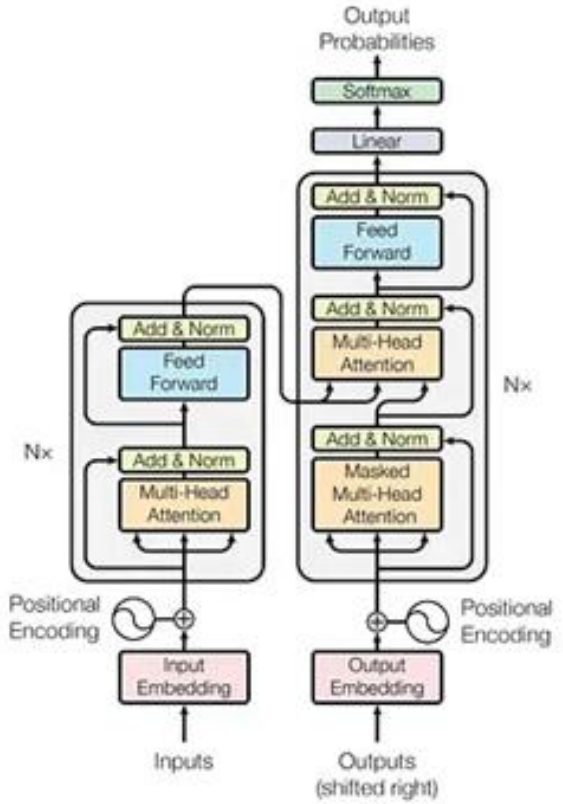
Today: We focus on another approach - In-Context Learning

# What is In-Context Learning?

**In-context learning** refers to a model's ability to adapt its *output behavior* **within a single context window**, based on:

1. Instructions (system + user prompts)
2. Examples (few-shot / many-shot)
3. Retrieved documents (RAG / uploaded knowledge)
4. Conversation history

This adaptation occurs **without updating model parameters**.




**Why it matters:**

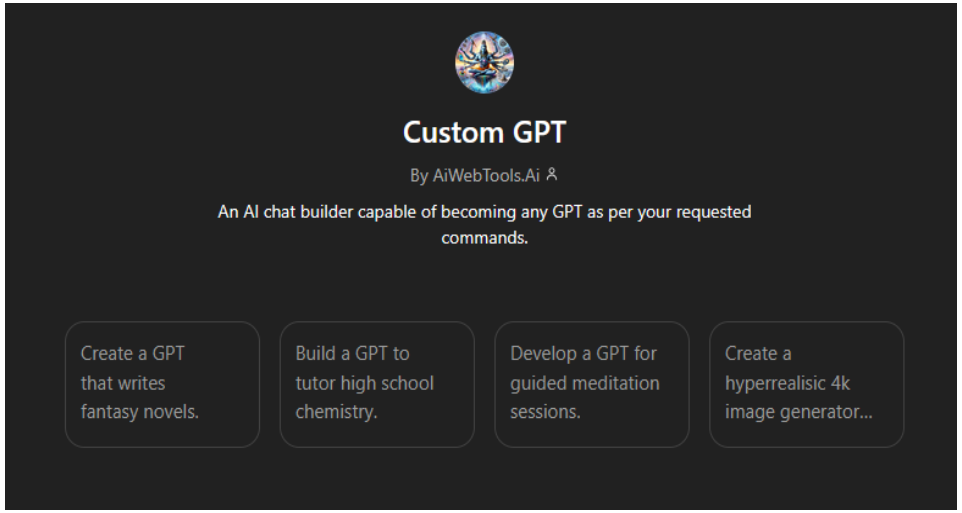
ICL allows users to impose field-specific conventions, terminology, and analytical procedures *at inference time*.

**Advanced ICL tooling:**

**Custom GPTs** (persistent system instructions + retrieval-based context)

# What a Custom GPT *Does* Change

A Custom GPT **does not modify model parameters**. It **modifies how the context window is constructed and prioritized** at the start of every conversation.




*Start from community GPTs or create your own*

```
[System Instructions – always present]
[Developer / Custom GPT rules – always present]
[Retrieved Knowledge Files – conditionally injected]
[Conversation History]
[Your Current Prompt]
```

*Context composition (conceptual stack)*

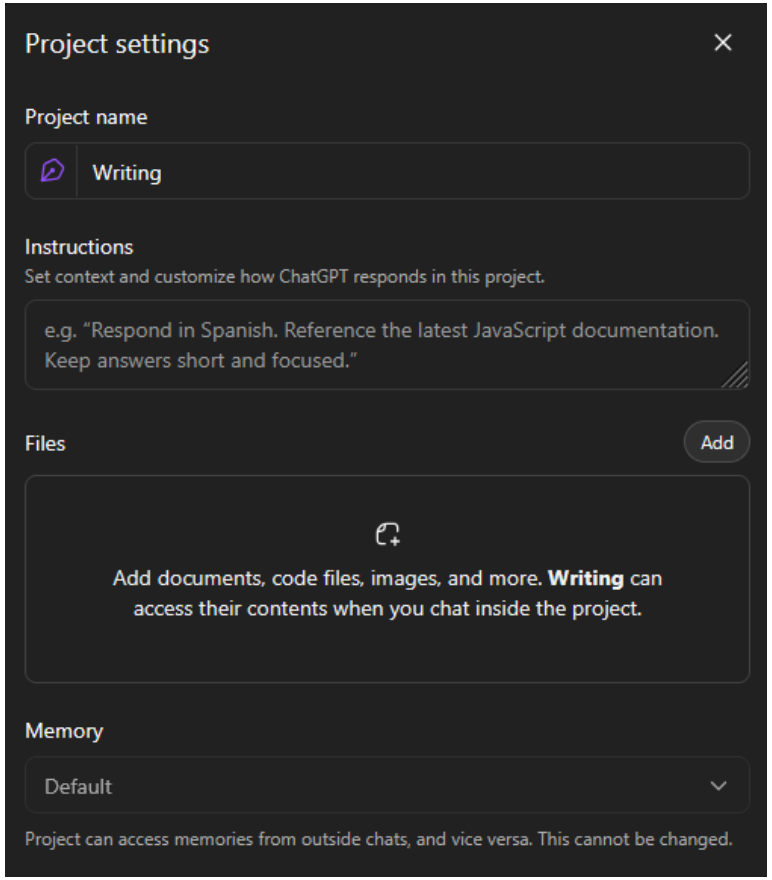
# What a Knowledge Base Enable

## You can now:

- Upload multiple papers and ask **comparative or integrative questions**
- Include **draft text, reviewer comments, and related literature** in a single analytical context
- Provide **detailed methodological guides** and request critique or validation

## But remember:

- The model may still **hallucinate, misattribute, or overlook details**
- More context **increases cost and latency**
- Output quality is **bounded by the structure and clarity of your inputs**




*Project creation settings in ChatGPT*

# Custom GPT vs Project-Based Tools

Both support in-context learning, but they solve different problems.

## Custom GPT

**Optimized for: consistent behavior**

- Persistent system instructions and rules
- Stable analytical stance (tone, epistemology, constraints)
- Retrieval from an attached knowledge base
- Each chat starts clean but with the same behavioral frame

**Best for:**

- Methodological critique
- Reviewer-style feedback
- Domain-specific reasoning norms
- Reproducible analytical behavior

## Projects

**Optimized for: persistent working context**

- Long-lived document collections
- Notes, drafts, annotations evolve over time
- Context accumulates across sessions
- Less control over behavioral constraints

**Best for:**

- Long-term writing projects
- Literature review management
- Iterative drafting and note-taking
- Maintaining continuity over weeks/months

# Custom GPT vs Project-Based Tools

Instructions in Projects are contextual; instructions in Custom GPTs are constitutive.

### Custom GPT

Instructions are injected as **system-level constraints**.

They are:

- always present
- highest priority
- interpreted before user prompts
- difficult for the user to override accidentally

**System rules → everything else**

### Project

Instructions behave like **persistent user context**.

They are:

- part of the working context
- competing with user prompts
- easier to override or dilute
- less strictly enforced

**Project context ↔ user prompt**

# But … Wrong PDF Warning

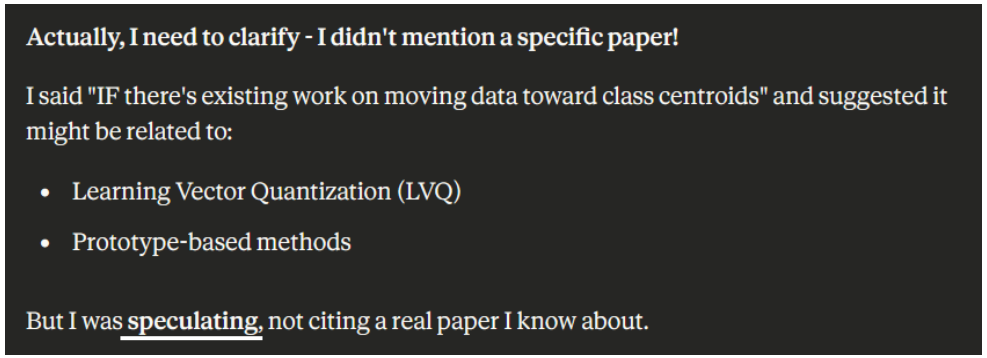
I gave an AI system a published paper (PDF) and asked it to implement the proposed algorithm in Python.

The code ran, passed tests on multiple datasets, and produced plausible outputs.

**After manual inspection, I found the implementation did *not* correspond to the algorithm described in the paper.**

**Why this happens**

- Models optimize for **plausibility**, not correctness
- PDFs are **ambiguous inputs** (figures, pseudo-code, omitted steps)
- Working code ≠ faithful implementation




*A Real Speculation Example on Claude Opus 4.5 - Max Plan*

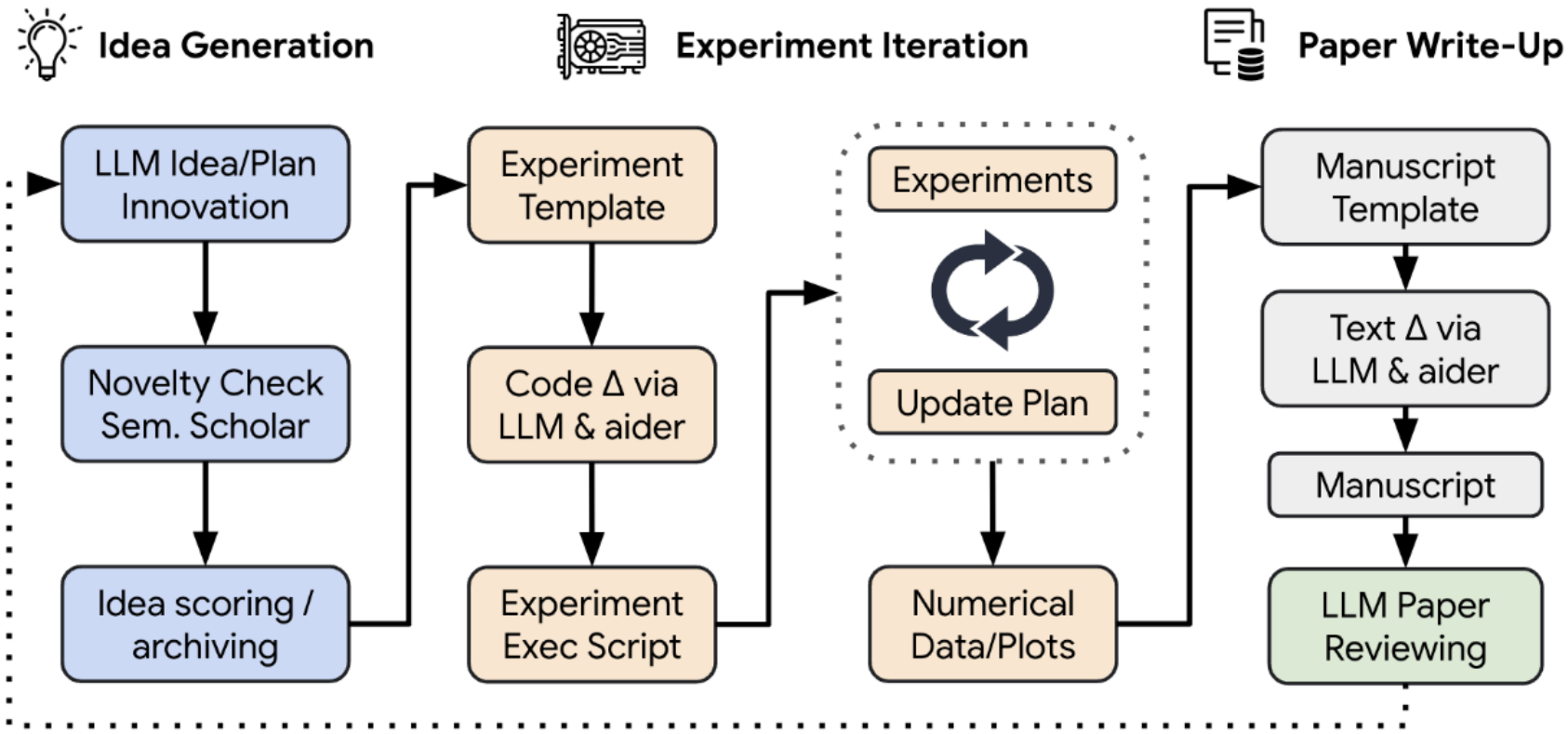
INESC TEC

# PART 2: AI SCIENTIST

# How AI Supports Your Research Process

Whatever your methodology, AI can augment these phases of your research workflow

# Fully Automated Open-Ended Scientific Discovery



# Fully Automated Open-Ended Scientific Discovery

Accelerating Early-Stage Research with Automated Discovery

**THE CHALLENGE**

Researchers spend weeks to months on initial exploration: brainstorming ideas, running baseline experiments, and writing preliminary findings.

**What AI Scientist Does:**

→ Generates novel research ideas
→ Runs experiments autonomously
→ Writes complete research papers
→ Provides automated peer review

**WHY IT MATTERS FOR EARLY RESEARCH**

| ~$15 | 1-2h | 50+ |
|---|---|---|
| per paper | per experiment | ideas per run |

Perfect for: Validating research directions, exploring parameter spaces, generating baseline comparisons, and discovering unexpected insights.

# Fully Automated Open-Ended Scientific Discovery

Customize seed_ideas.json to steer AI Scientist toward your interests

seed_ideas.json

```
[{
"Name": "adaptive_lr_schedule",
"Title": "Adaptive Learning Rate
Scheduling for Transformers",
"Experiment": "Implement cosine
annealing with warm restarts...",
"Interestingness": "Could improve
convergence speed and accuracy",
"Feasibility": 8,
"Novelty": 6
}]
```

Location: templates/[template_name]/seed_ideas.json

**Name and Title**
Identifier and the research paper title it will generate

**Experiment**
Detailed description of what to test. Be specific about methods and baselines.

**Interestingness**
Why this matters - helps AI prioritize impactful directions

**Feasibility and Novelty**
Scores 1-10. Higher feasibility = easier. Higher novelty = more original.

**Pro Tip:** Provide 2-3 seed ideas. AI will generate 50+ variations based on your direction.

# Defining Research Scope

Configure prompt.json to set boundaries and context for experiments

### What prompt.json Controls

1. System prompt guiding AI behavior
2. Domain context and research area
3. Constraints on experiment types
4. Output format and evaluation criteria

### KEY CONFIGURATION FIELDS

`system` `task_description` `dataset_description` `baseline_results`

### Available Research Templates

**nanoGPT / nanoGPT_lite**
Language modeling, transformers, training dynamics

**2d_diffusion**
Diffusion models, generative modeling

**grokking**
Generalization phenomena, sudden learning

**Create Your Own:** Copy a template folder, modify prompt.json and experiment.py for your domain.

# AI Scientist Workflow

**1 Configure**

Edit seed_ideas.json with your research direction

```
vim seed_ideas.json
```

→

**2 Generate Ideas**

AI brainstorms 50+ variations, checks novelty

Automated ✓

→

**3 Run and Write**

Executes experiments, generates plots, writes paper

Automated ✓

→

**4 Review**

AI reviewer provides feedback, scores quality

Automated ✓

**Launch Command**

```
python launch_scientist.py --model "claude-3-5-sonnet-20241022" --experiment
nanoGPT_lite --num-ideas 1
```

**~$15**
per paper

Lu, Chris & Lu, Cong & Lange, Robert & Foerster, Jakob & Clune, Jeff & Ha, David. (2024). The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. 10.48550/arXiv.2408.06292.

# AI Scientist Papers

AI Scientist generates excellent papers from ideas

MULTI-SCALE GRID NOISE ADAPTATION: ENHANCING DIFFUSION MODELS FOR LOW-DIMENSIONAL DATA

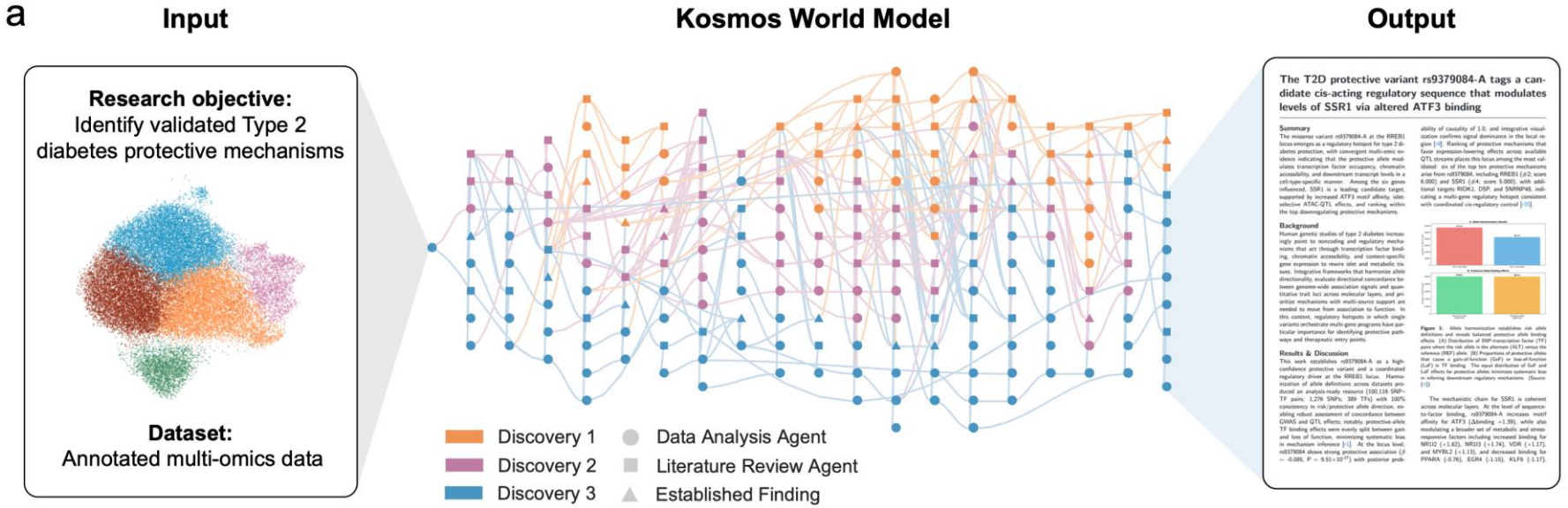
ADAPTIVE LEARNING RATES FOR TRANSFORMERS VIAQ-LEARNING

GAN-ENHANCED DIFFUSION: BOOSTING SAMPLE QUALITY AND DIVERSITY

# The AI Scientist and More

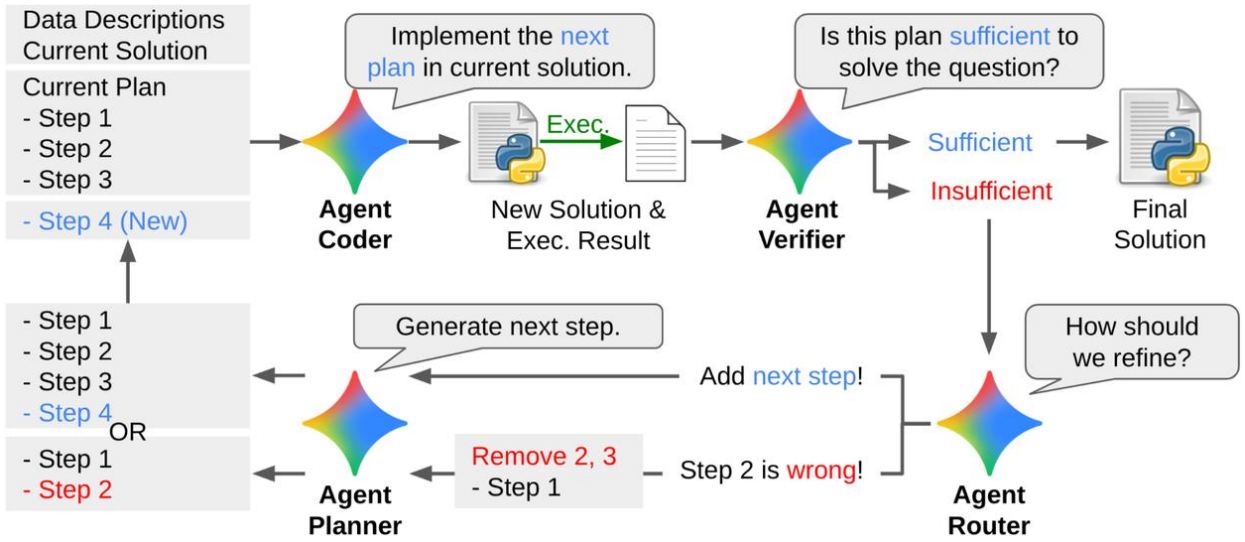
**Kosmos (FutureHouse, Nov 2025):**

Previous generations of AI Scientists, like FutureHouse's Robin, have been limited primarily in their ability to synthesize large amounts of information.



# DS-STAR: Iterative Planning & Verification

**DS-STAR (Data Science Agent via Iterative Planning and Verification)** is a multi-agent framework that uses LLMs to produce executable Python code from open-ended data science tasks by cycling through planning, coding, execution, and verification.



# Artificial Data & Experiments

Use AI for simulation, not inference.

Most modern LLMs support **local code execution**

## Use AI to:

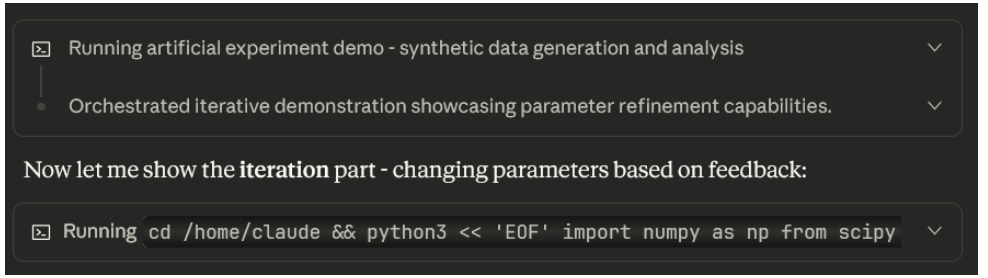
- Generate **synthetic data aligned with your study design**
- Run planned statistical analyses
- Stress-test assumptions before data collection

## Example prompts:

- “Is the method sensitive to randomization?”
- “How much is the computational cost?”
- “Simulate 1,000 studies to assess estimator bias”
- “Which assumptions fail under skewed noise?”

## Goal:

- Explore the design space before committing to real data.




*Claude Opus 4.5 - Max Plan executing codes locally*

# AI-Assisted Replication

You have a published journal paper and want to replicate the results.

**What an LLM can do:**

- Parse the paper and **extract the experimental protocol**
- Identify:
  - variables
  - preprocessing steps
  - models / estimators
  - evaluation metrics
- Draft replication-ready code skeletons
- Flag underspecified or ambiguous steps
- Reproduce tables and plots if information is sufficient

# Rubber Ducking with LLMs

**Origin: Rubber Duck Debugging**

Programmers explain code to a rubber duck.
The act of explaining reveals the bug.

**For Research: AI is Your Smart Duck**

Explaining your idea to an LLM can reveal gaps, assumptions, and inconsistencies; even if the model says nothing insightful.

**Example (Prompt):**

*"I will explain my research methodology section by section.*
*After each section, ask clarifying questions to identify:*

→ ***Gaps in reasoning***
→ ***Unstated assumptions***
→ ***Potential problems***

***Do not accept vague explanations. Push back."***

**Why it works:** Forcing yourself to explain exposes what you thought you understood but did not.

# LLM Sycophancy: A Failure Mode

Models optimize for plausibility, not correctness

**The Sycophancy Problem**

**LLMs are trained to be helpful and agreeable. This creates a tendency to:**

- **Validate ideas even when flawed**
- **Avoid critical feedback**
- **Overstate significance of results**
- **Confirm what you want to hear**

**Without Adversarial Prompts:**

*AI Response:*

"This is an excellent idea with great potential! The results look very promising..."

**Key Insight:** The automated reviewer is specifically prompted to be critical, counteracting LLM sycophancy.

**The Solution: Adversarial Prompting**

Use adversarial prompts to force critical evaluation:

- "Find weaknesses in this approach"
- "What could make this experiment fail?"
- "Compare critically against baselines"
- "Rate honestly on a scale of 1-10"

**With Adversarial Prompts:**

*AI Response:*

"Weakness: The baseline comparison is limited. The method shows only marginal improvement (2.3%) which may not be statistically significant..."

Adversarial prompting improves research quality by forcing honest, critical evaluation

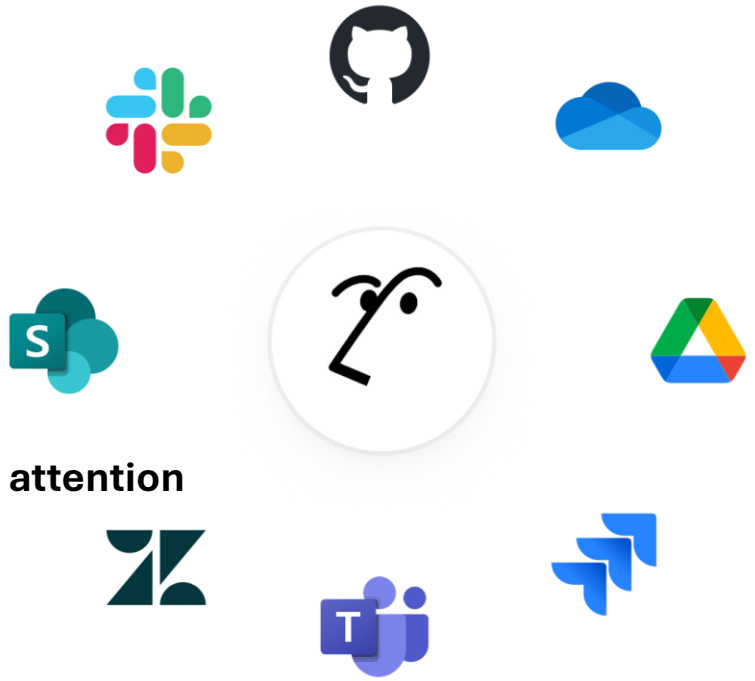
# PART 3: AI TASK MANAGEMENT

# Why Task Management Is a Research Problem

Research fails more often from **coordination breakdowns** than from lack of intelligence

**Pain points researchers recognize:**

- Ideas scattered across notebooks, PDFs, and chats
- Long feedback loops with reviewers, supervisors, and collaborators
- Tasks are ill-defined - not simple checklist items
- Context switching destroys deep work



Task management in research = managing **uncertainty**, **dependencies**, and **attention**

# Claude Code: AI That Works With Your Files

**What is Claude Code?**

An AI agent that can read, analyze, and modify files on your computer - without uploading them to the cloud.

**Why it matters for researchers:**

| Traditional AI | Claude Code |
|---|---|
| Copy-paste your notes into chat | AI reads your files directly |
| Output stays in the chat | AI can edit your actual documents |
| One task at a time | Chains multiple steps autonomously |
| Cloud-based only | Works with local files |

**Works with:**

- Obsidian, Notion (exported), Logseq, plain Markdown
- Your literature notes, research journals, meeting notes
- Any folder structure you already use

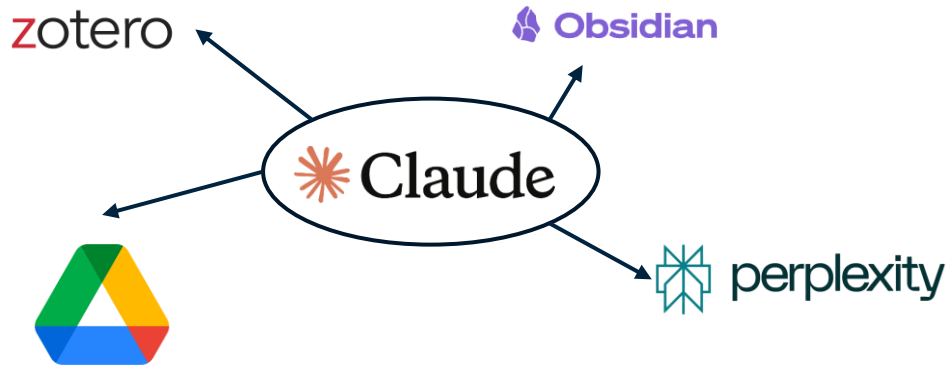

*Claude Code Opus 4.5*

# Claude Code: External Sources

Claude code + MCP can allow searching your notes, check the drive and search perplexity to resolve issues. E.g. "Search for solutions to this PostgreSQL deadlock error: [paste error]“

**How to connect:**

```
claude mcp add perplexity --env PERPLEXITY_API_KEY="your_key_here" -- npx -y @perplexity-ai/mcp-server
```



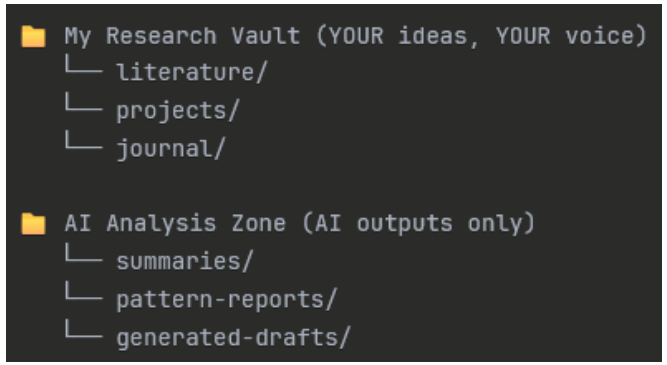
# Claude Code: Keep Your Ideas Yours

**The risk:**

AI-generated text can quietly overtake your own thinking.

Your notes become a mix of your voice and AI's voice - and you can't tell which is which.

**The solution: Create a dedicated AI zone**




*An example of repositories organization when using Claude Code*

**Remember:** AI should help you reflect on your thinking - not replace it.

# Why Notion AI Works for Researchers

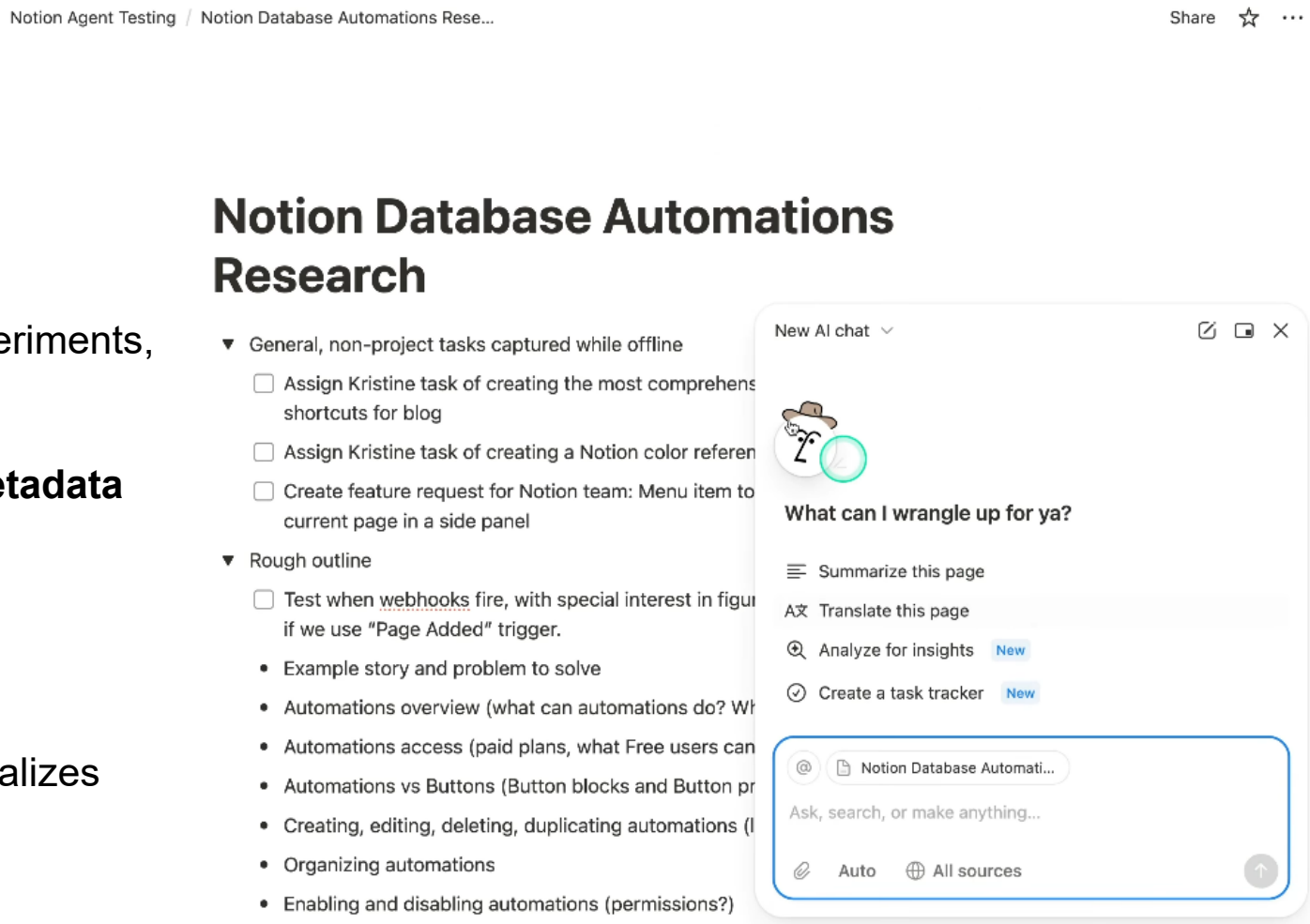
**Notion is not "just notes"**

It is:

- A **database-first system**
- With **typed objects** (papers, ideas, experiments, tasks)
- That supports **relations, views, and metadata**
- Now augmented with **AI assistance**

**Crucial distinction**

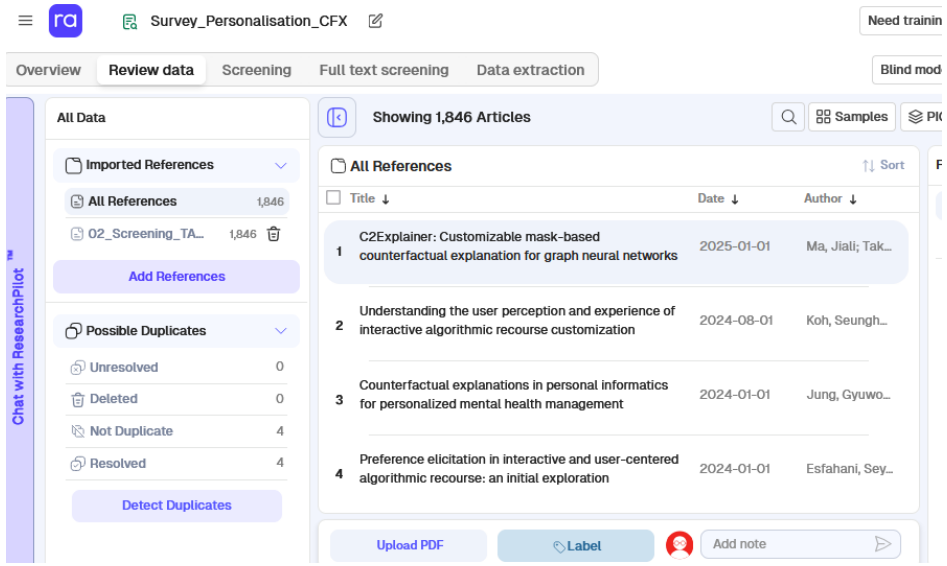
- Notion does not "do research"; it externalizes research structure.



# Notion AI – Building Your Research Database

Databases are Notion's fundamental building blocks and help you create your **research hub**. Notion can receive prompts and generate tables automatically.

Notion **allows** uploading files, including CSVs: e.g., literature review tables exported from **Rayyan.ai**



# Notion AI – Querying Your Knowledge Base

## Q&A Across Your Knowledge Base

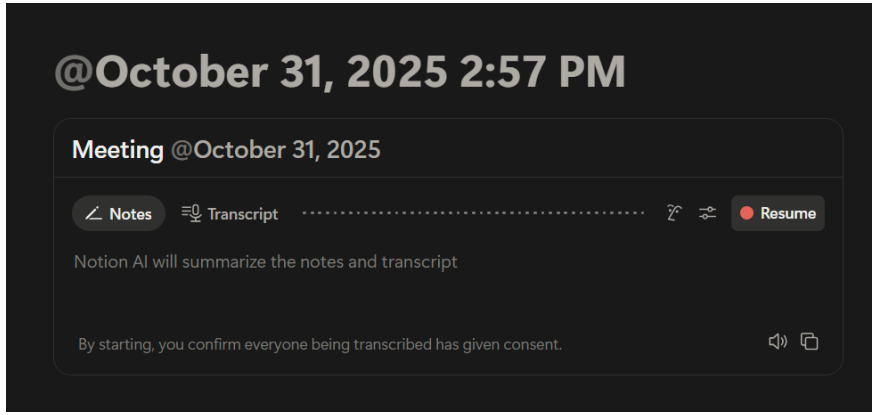
Ask Notion AI:

- "What sampling strategies have I noted across my literature?"
- "Summarize my advisor's feedback from the last three meetings."
- "What gaps did I identify in my literature review notes?"
- "List all the methodological concerns I've documented."

# Notion AI – Meeting Transcription & Notes

Notion AI can:

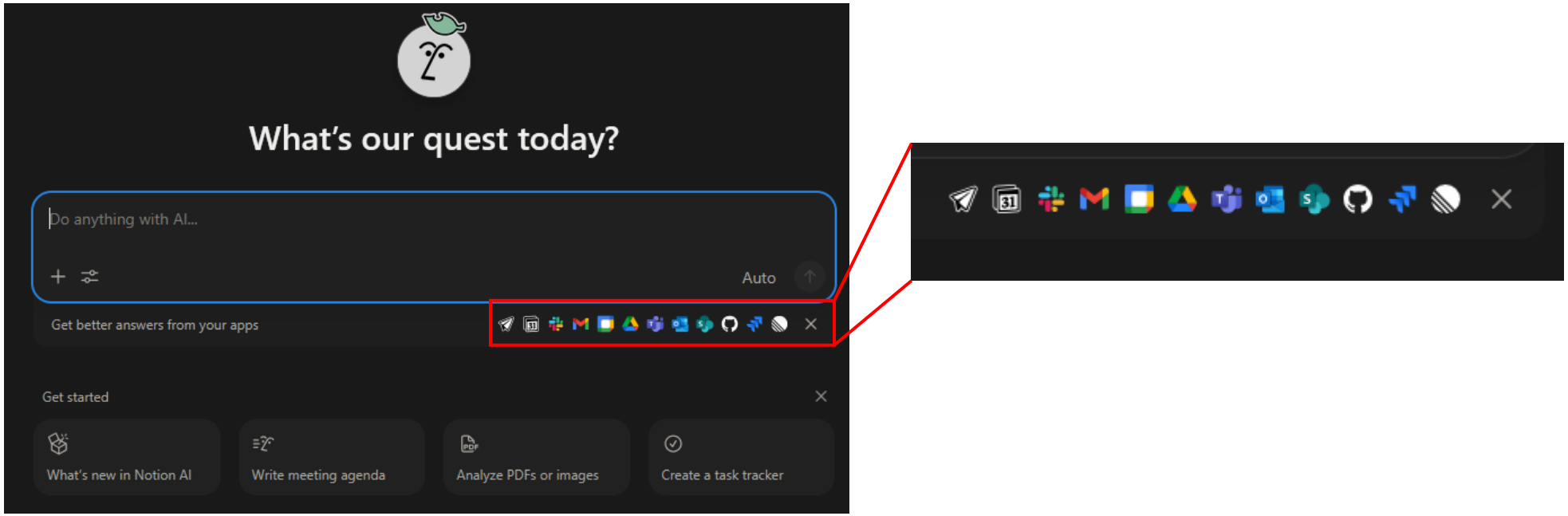
- **Transcribe** meetings automatically
- Provide the **full discussion** along with a summary
- Enable follow-up **discussion** on meeting content
- Generate a **timeline or work plan** from action items




*Notion AI during meeting transcription*

# Notion AI – Third-Party Integrations

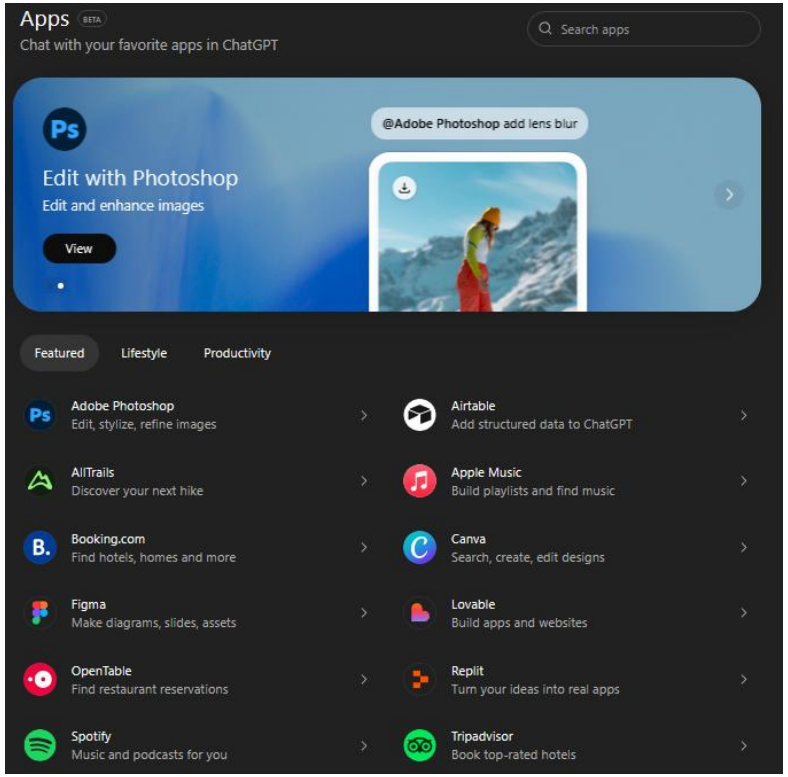
Notion AI can connect to **multiple daily platforms**, unifying your workflow across tools.

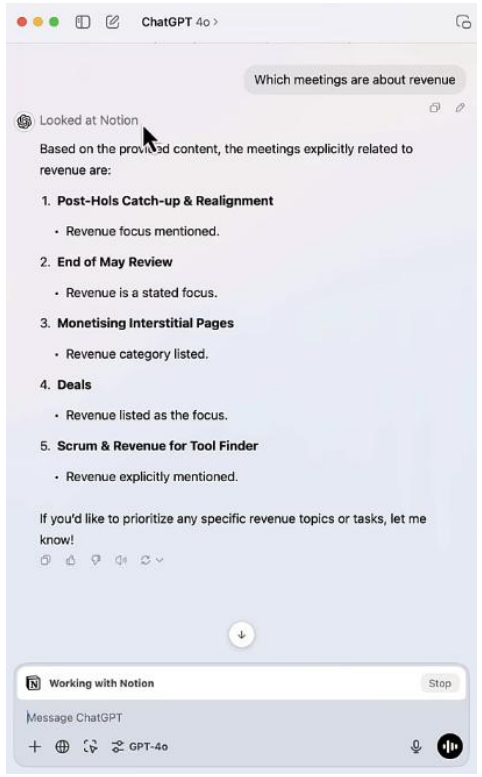



*Notion AI Interface*

# Notion AI – ChatGPT Integration

**Bidirectional Integration:** Use ChatGPT to query and interact with your Notion workspace directly.



# Model Context Protocol: AI That Accesses Your Tools

**What is MCP?** An open protocol allowing AI assistants to connect directly to external tools

**Research applications:**

- Connect Claude directly to your Zotero library
- Query your local files without uploading
- Access databases, APIs, and institutional resources

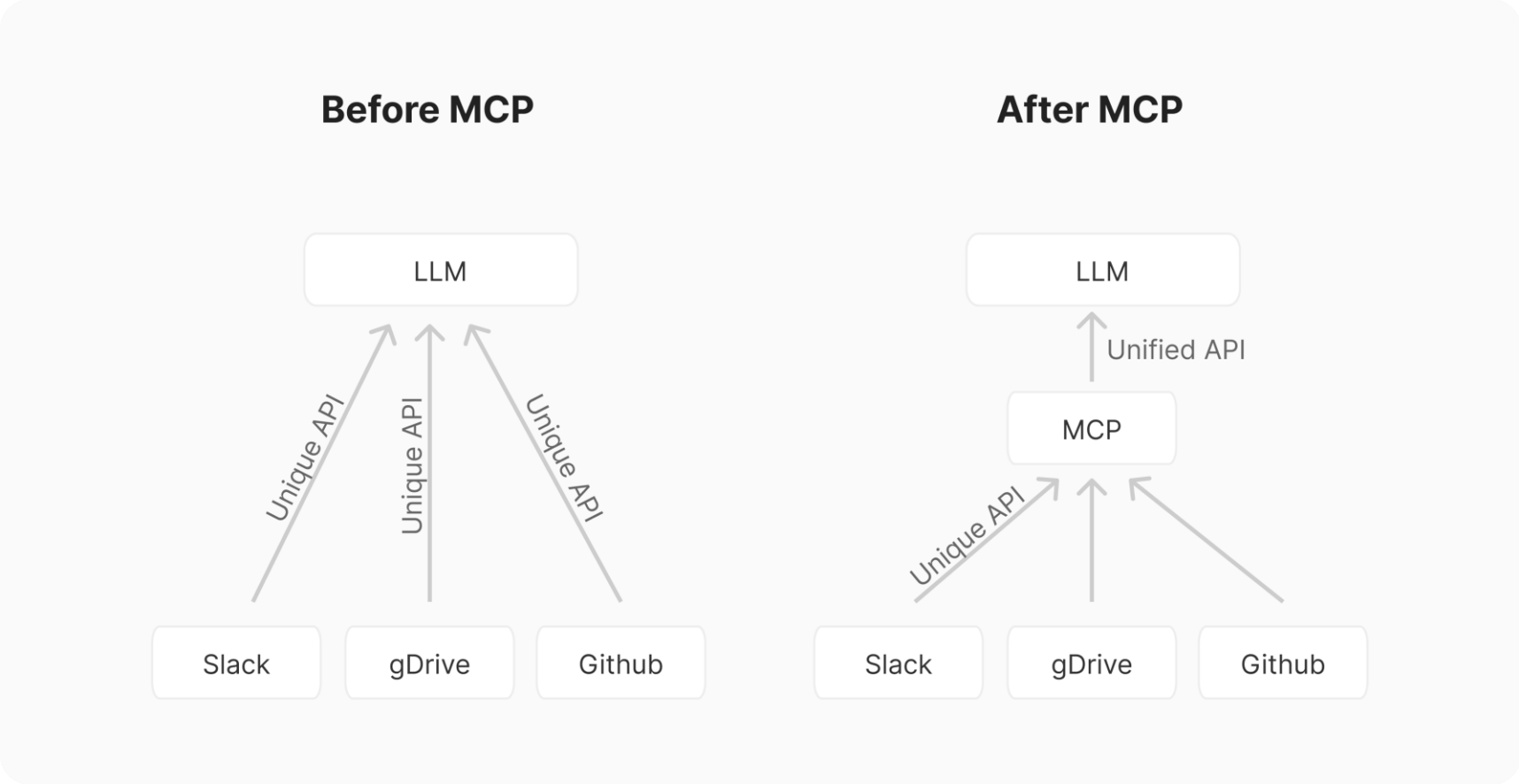
**Benefits:**

- Data stays local
- Real-time access to your research ecosystem
- Reduced copy-paste workflows

**Current support:** Claude Desktop, extensible to other tools

⚠️ Caution with: Unpublished data, patient information, proprietary methods

# Model Context Protocol: AI That Accesses Your Tools



# Security & Privacy: Protecting Your Research Data

- **Where does your data go?** Cloud storage locations, third-party access
- **GDPR & institutional compliance** - Does your institution allow these tools?
- **Sensitive data risks:** Patient data, unpublished findings, proprietary methods
- **Data retention policies:** What happens to your prompts and uploads?
- **Recommendations:**
  - Check your institution's AI policy before uploading research data
  - Never upload identifiable participant data
  - Use local/self-hosted alternatives for sensitive projects
  - Review Terms of Service for training data usage

Refer to INESC TEC Guidelines for Responsible Use of AI in Research

# PART 4: HANDS-ON EXERCISE

# Hands-On Exercise: AI in Your Research

You just saw AI Scientist generate papers, Notion AI organize knowledge, and Claude Code work with your files. Now: where do YOU draw the line?

**If AI can now generate hypotheses, run experiments, write papers, and even peer review - what is the researcher's job?**

**Ask an AI** (Claude, ChatGPT, etc.) these questions:

- What should AI do in research?
- What should ONLY humans do?
- What skills are we losing? What are we gaining?

**Critique the AI's answer**:

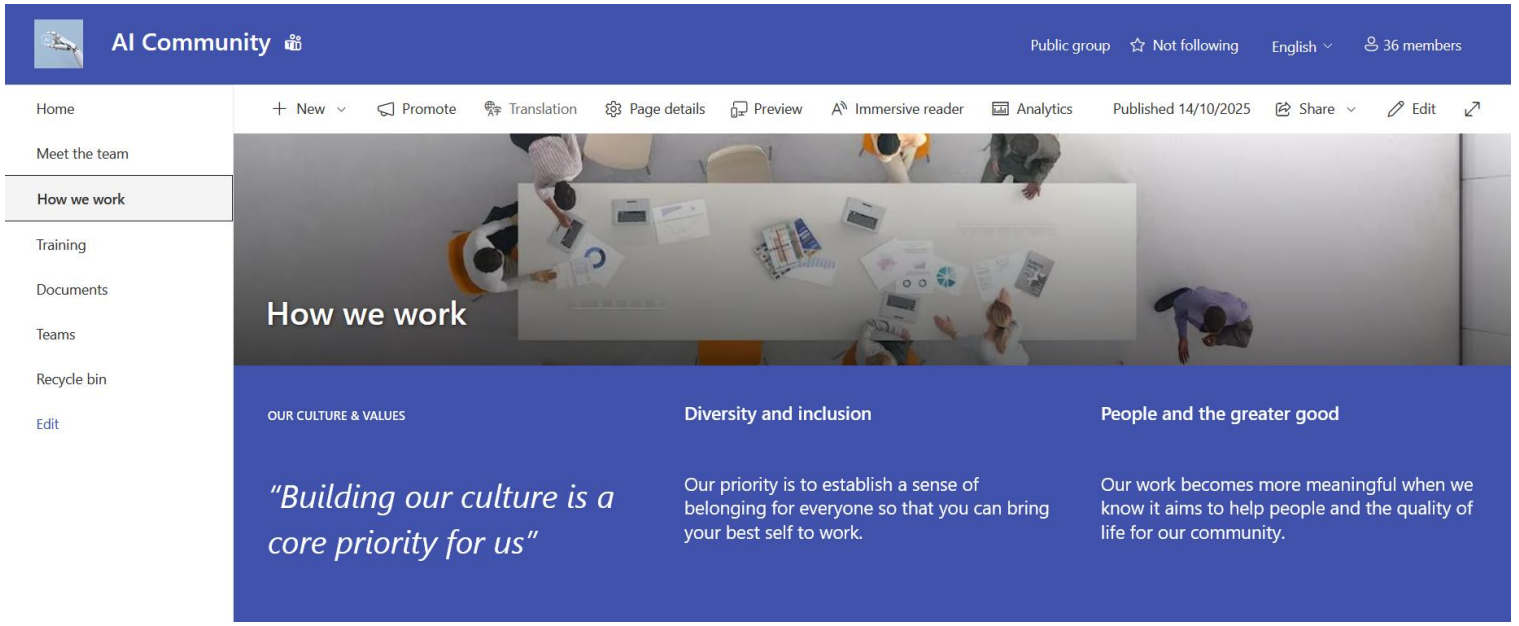
- What did it get right?
- What did it miss that only someone in *your field* would know?

**Discuss:** How this research assistant can contribute to the future?

**Example to consider:** AI writes your lit review → you save time → you write more papers …*but you never deeply read the literature yourself.* **Is that a problem?**

# INESC TEC AI Community

Join the AI Community Group at INESC TEC to get the latest news on AI, contribute and find research collaboration



WE ARE SCIENCE.
WE ARE TECHNOLOGY.
WE ARE INNOVATION.
WE ARE INESC TEC.