

[Skip to main content](#)

ClaudeAI X

Search in r/ClaudeAI

[Log In](#)

...

 r/ClaudeAI • 5mo ago
bbleimschein

...

I Built an AI-Powered Research Pipeline with Claude Agents That Turns Scientific Papers into Blog Posts

TL;DR

With Claude Code and two custom Claude Agents, I built an end-to-end editorial system that discovers, analyzes, and turns scientific papers into long-form posts. In ~3 weeks it ingested **116** papers, deeply analyzed **129**, flagged **35+** potential breakthroughs ($\geq 8/10$ quality), and auto-published **~15** posts to my Next.js blog. **Here is today's feature:** [The Mirror of Intelligence: How AI Systems Are Learning to See the World Through Human Eyes](#)



Landing page of Futurelab.Studio

The Problem I Solved

Every day, thousands of papers drop. Somewhere in that torrent are the few that change how we think—but they're buried in jargon, scattered across databases, and rarely reach practitioners. I wanted Claude Agents to do what a human editorial team would: scan broadly, assess rigor, connect threads, and tell a coherent story—then send it right to my inbox.

I asked myself: **What if Claude Agents could build an entire editorial pipeline - from research discovery to published content?**

What I Built: A Complete AI Editorial System

The Architecture: Three Claude Agents, Two Databases, One Mission

1. Research Paper Analyst Agent

- Fetches papers from 6 sources (arXiv, OpenAlex, PubMed Central, etc.)
- Analyzes each paper with structured assessments: quality ratings (1-10), breakthrough potential, methodology review
- Stores analysis in local PostgreSQL database with pgvector embeddings
- Cross-references existing research to avoid duplicates and find connections

2. Science Storyteller Agent

- Reads analyst's assessments from local database
- Identifies themes across multiple papers
- Writes ~3000-word blog posts with IEEE citations

[Skip to main content](#)[Log In](#)

The Technical Stack:

- **Open-Source MCP Server** ([Github](#)) - Standardized tool APIs (incl. rate limits, etc.)
- **Local Analysis Database & MCP server** - PostgreSQL with pgvector for semantic search
- **Production Blog** - Next.js + PostgreSQL on [futurelab.studio/blog](#)
- **Claude Agent Configs** - Custom `.claude/agents/` definitions
- **Observability Dashboard** - Real-time monitoring of all operations

The Magic Numbers:

- **129 papers** ingested across disciplines
- **111 papers** analyzed with quality assessments (some papers don't have content available)
- **35+ breakthroughs** identified (8+ quality rating)
- **~15 blog posts** auto-published
- **2 Claude Agents** working in harmony (Analyst + Story Teller)

How I Built It With Claude Code

This started in the most human way: I was dragging PDFs into a chat window because I liked Balaji's line about "wisdom in the latest papers and old texts"—and I don't love academic prose. Manual uploads got me quick insights, but not a process. The first breakthrough was writing a **Scientific Papers MCP server** so Claude could fetch, parse, and summarize without me screening an uploading anything.

From there, I pushed the work into **Claude Code** with two custom agents. The Analyst's job: "summarize the abstract," and "act like a reviewer." I gave it a schema: extract the claim in <120 words, justify a 1–10 quality score with concrete method checks (data, ablations, baselines) and estimate breakthrough potential. Then, save the analysis to my local Postgres (+ pgvector). That schema did two things: it made outputs comparable day-to-day, and it made the database useful for theme-finding.

The Storyteller began as a drafting helper and ended up as a finisher. It reads the top N analyses from today, clusters them for a unifying theme ("Are these three really about 'AI systems learning to think'?"'), and writes a long-form post that cites sources and links to prior posts to build narrative continuity. The last bit was wiring a small publish endpoint in my Next.js blog. Once that existed, the system flipped from "assistive" to "autonomous." (Currently still with a bit of human oversight.)

I plan like a product manager but implement like a tinkerer. My lightweight PM routine (I swear by the [Bmad-Method](#)) gives me a PRD, an architecture outline and user stories, but most importantly establishes the practice of test driven development. The process encompasses everything from business analysts, product managers, architects, and so on. After that it's fast iterations: implementing with Claude Code and getting the whole thing deployed. The result isn't a monolith; it's a set of small, well-behaved components that Claude can orchestrate.

The Agent Pipeline at Work

Step 1: Research Analyst in Action

[Skip to main content](#)[Log In](#)**breakthrough ratings).**

- I'll use the research-paper-analyst agent to analyze the latest AI papers from arXiv and store them in the database with detailed methodology and breakthrough ratings.
- **research-paper-analyst**(Analyze latest AI papers)
 - └ Done (30 tool uses • 72.7k tokens • 5m 6.9s)
- The research-paper-analyst agent has completed analyzing the latest AI papers from arXiv. The agent fetched and analyzed 20 recent papers, storing both the paper metadata and detailed analysis in the database.

Key Highlights:

The agent identified several high-impact papers with breakthrough potential ratings of 8-9, including:

1. **LiveMCP: Temporal Reasoning for Long-Context Live Streaming Interactions** (Breakthrough: 9/10)
 - Addresses critical live streaming challenges with novel temporal architecture
2. **Towards System 2 Reasoning in LLMs: Learning How to Think With Meta Chain-of-Thought** (Breakthrough: 8/10)
 - Introduces meta-reasoning capabilities for improved problem-solving
3. **Evaluating and Enhancing LLMs for Multi-turn Text-to-SQL with Multiple Data Sources** (Breakthrough: 8/10)
 - Advances multi-database SQL generation capabilities
4. **MAViS: Modular Autonomous Virtuoso Software Generation Framework** (Breakthrough: 8/10)
 - Novel approach to autonomous software generation

All papers have been stored with comprehensive methodology descriptions, quality ratings, and breakthrough assessments in the database for future reference and analysis.

The research paper analyst agent in action in the CLI

Step 2: Science Storyteller Creating Content

[Skip to main content](#)

X

[Log In](#)

- I'll use the science-storyteller agent to fetch the recently analyzed papers, craft an engaging blog post focusing on the highest-rated discoveries, and publish it to the futurelab blog.

- **science-storyteller**(Create and publish research blog post)
 - ↳ Done (11 tool uses · 53.2k tokens · 6m 5.0s)

- Perfect! The science-storyteller agent has successfully created and published an engaging blog post to futurelab.studio.

Published Post: "The Mirror of Intelligence: How AI Systems Are Learning to See the World Through Human Eyes"

The agent focused on the highest-rated papers from today's analysis, particularly featuring the groundbreaking Brain-AI Convergence research (Quality 9, Breakthrough 9) as the centerpiece. The 3,100-word post weaves together four major breakthroughs into a compelling narrative about AI systems naturally developing human-like capabilities.

The post has been published as a draft on futurelab.studio and is ready for your review and potential publication. The storytelling approach makes complex research accessible while maintaining scientific accuracy, perfect for engaging your tech-savvy audience.

> ■

bene@Benedicts-MacBook-Pro:~/Documents/Dev-Source/scientific_papers [Sonnet 4]
►► accept edits on (shift+tab to cycle)

o

The story teller agent in action, pulling the research and writing the blog post

Step 3: Scientific papers repository and observability platform (local)

[Skip to main content](#)[Log In](#)

Latest Research Discoveries

Explore 129 scientific papers with AI-powered analysis and quality ratings

129
Total Papers

111
Analyzed

7.6
Avg Quality

29
This Week

Filter by Source:

- [All Sources](#)
- [ARXIV 122](#)
- [BIORXIV 5](#)
- [DEMO 1](#)
- [PMC 1](#)

ARXIV Q: 8/10 B: 7/10

CMPHysBench: A Benchmark for Evaluating Large Language Models in...

Benteng Chen, Binzhao Luo, Changyong He, Di Zhang,...

August 25, 2025

Why It Matters:
CMPHysBench addresses a critical need in the intersection of AI and scientific research by providing a standardized evaluation framework for assessing larg...

[Read Full Analysis →](#) 75% confidence

ARXIV Q: 7/10 B: 7/10

Type-Compliant Adaptation Cascades: Adapting Programmatic LM Workflows t...

Chu-Cheng Lin, Dalyi Peng, Eugene Ie, Ming Zhang,...

August 25, 2025

Why It Matters:
This research addresses a critical challenge in deploying language models at scale: how to adapt complex programmatic workflows to diverse data...

[Read Full Analysis →](#) 68% confidence

ARXIV Q: 8/10 B: 8/10

Efficient Computation of Blackwell Optimal Policies using Rational Functions

Dibyangshu Mukherjee, Shivaram Kalyanakrishnan

August 25, 2025

Why It Matters:
This research addresses a fundamental computational challenge in decision theory and reinforcement learning: efficiently computing Blackwell optimal...

[Read Full Analysis →](#) 55% confidence

ARXIV Q: 6/10 B: 7/10

Hermes 4 Technical Report

Chen Guang, Dakota Mahan, Jai Suphavadeepasit, Jeffrey Quesnelle,...

August 25, 2025

Why It Matters:
Hermes 4 represents a significant advance in open-source large language models by combining structured multi-turn reasoning capabilities with broad...

[Read Full Analysis →](#) 65% confidence

ARXIV Q: 9/10 B: 9/10

Disentangling the Factors of Convergence between Brains and Computer Vision...

Camille Couprie, Huy V. Vo, Jean-Rémi King, Joséphine...

August 25, 2025

Why It Matters:
This research provides fundamental insights into one of the most intriguing phenomena in modern AI: why artificial neural networks trained on natural images...

[Read Full Analysis →](#) 94% confidence

ARXIV Q: 9/10 B: 8/10

SafeBimanual: Diffusion-based Trajectory Optimization for Safe Bimanual...

Haoyuan Deng, Qianzhen Wang, Wenkai Guo, Zhenyu Wu,...

August 25, 2025

Why It Matters:
SafeBimanual addresses a critical gap in robotic manipulation by introducing safety constraints to diffusion-based bimanual manipulation policies. This...

[Read Full Analysis →](#) 92% confidence

Overview of the latest analysed papers

Step 4: Published, final blog post

[Skip to main content](#)[Log In](#)

Research for the curious mind

[All Posts](#)

The Mirror of Intelligence: How AI Systems Are Learning to See the World Through Human Eyes

August 26, 2025

In the vast landscape of artificial intelligence research, few discoveries have been as profound and unexpected as this: AI systems trained on natural images are spontaneously developing representations that mirror the human brain. Not metaphorically, not approximately—but with such striking similarity that neuroscientists can now predict brain activity from AI model responses and vice versa. This isn't the result of explicit programming or neural network architectures designed to mimic biology. It's the natural consequence of artificial minds learning to see.

Today's breakthrough research reveals why this convergence happens, how it unfolds during training, and what it means for the future of both artificial intelligence and our understanding of human cognition. Combined with revolutionary advances in robot safety systems, comprehensive AI evaluation frameworks, and adaptive programming paradigms, August 26 represents a pivotal moment when AI research achieved new depths of understanding about intelligence itself—both artificial and biological.

But this story begins with a question that has puzzled scientists since the earliest days of computer vision: Why do artificial neural networks, designed purely for practical tasks like image recognition, end up organizing visual information in ways that so closely mirror the human brain?

The Mysterious Convergence: When Artificial Minds Echo Human Vision

For years, researchers have observed a remarkable phenomenon: AI models trained on image classification tasks develop internal representations that correlate surprisingly well with neural activity in human visual cortex. This brain-AI similarity has been documented across multiple model architectures and training approaches, yet the fundamental mechanisms driving this convergence remained mysterious.

The breakthrough research by an international team led by scientists from prestigious institutions across Europe and the United States provides the first systematic investigation into what drives this extraordinary phenomenon [1]. Their findings don't just explain why brain-AI convergence occurs—they reveal it follows predictable patterns that mirror human brain development itself.

Final product, right in my blog on futurelab.studio/blog (and in my inbox)

What I'd love feedback on

I'd love your feedback on the quality of the posts and if you leave me an upvote and a comment with a topic, I'd be happy to run my agents against it and publish a post!

Cheers!

PS: Example prompt (my research analyst)

You are a senior research analyst with 25+ years of experience in identifying and evaluating breakthroughs in scientific research. Your primary mission is to analyze research papers and create comprehensive, high-quality summaries that highlight the most significant findings and their implications. For every paper you analyze, you must provide a structured summary containing exactly these five sections:

****Why It Matters**:** A compelling 2-3 sentence explanation of why this research is significant to the field.

[Skip to main content](#)[Log In](#)

****Methodology**:** A brief but informative overview of the research approach, experimental design, data sources, and analysis methods.

****Quality Rating**:** Assign a score from 1–10 based on:

- Methodological rigor and experimental design (3 points)
- Statistical validity and sample size adequacy (2 points)
- Novelty and originality of approach (2 points)
- Clarity of presentation and reproducibility (2 points)
- Peer review status and publication venue (1 point)

Provide the numerical score followed by a 1–2 sentence justification.

****Breakthrough Potential**:** Rate from 1–10 the likelihood this research will lead to significant advances in its field.

- Paradigm-shifting potential
- Practical applicability
- Foundation for future research
- Cross-disciplinary impact

Provide the numerical score with a brief explanation of your reasoning.

Your analysis standards:

- Prioritize papers with genuine novelty over incremental improvements
- Be skeptical of extraordinary claims and look for robust evidence
- Consider both immediate and long-term implications
- Identify potential limitations or areas needing further validation
- Maintain objectivity while recognizing transformative potential

If a paper lacks sufficient detail for proper analysis, request the full text or additional information.

Your summaries should be comprehensive enough for researchers to understand the paper's significance without needing to read the full text.

Simple 2-Step Workflow

Step 1: Find and Analyze Papers

****Discover research**** using scientific papers MCP tools:

- `mcp_scientific_papers_fetch_latest(source="arxiv", category="cs.AI", count=20)` – Get latest papers
- `mcp_scientific_papers_search_papers(source="arxiv", query="your topic", count=10)` – Search by topic
- `mcp_scientific_papers_fetch_content(source="arxiv", paper_id="2305.11176")` – Get full content

****IMPORTANT – Avoid Duplicates**:** Before analyzing any paper, check if it already exists in the database.

- Use `mcp_local_research_blog_management_search_papers_text(query="paper_title_keywords", limit=5)` to search by title and keywords
- Use `mcp_local_research_blog_management_search_papers_text(query="arxiv_id", limit=5)` to check by arXiv ID
- Skip papers that are already in the database to avoid duplicates
- Focus on genuinely new papers that haven't been previously analyzed

****Analyze each promising NEW paper**** with your 5-section structure:

- **Why It Matters**: 5–10 sentences on significance and real-world impact
- **Findings**: Clear summary of key discoveries and results
- **Methodology**: Overview of research approach and methods
- **Quality Rating**: Score 1–10 based on rigor, validity, novelty, clarity, and peer review status
- **Breakthrough Potential**: Score 1–10 for likelihood of significant field advances in 5–10 years

Step 2: Store in Database

****Save each analyzed paper**** using:

```

```
First store the paper metadata
```

[Skip to main content](#)[Log In](#)

```

 title="Paper title",
 abstract="Paper abstract",
 authors=["Author 1", "Author 2"],
 publication_date="2025-01-01",
 url="https://arxiv.org/abs/2305.11176"
)

Then store your detailed analysis
mcp_local_research_blog_management_store_analysis(
 paper_id=paper_result["paper_id"],
 why_it_matters="Your 'Why It Matters' section content",
 findings="Your 'Findings' section content",
 methodology="Your 'Methodology' section content",
 quality_rating=8,
 quality_justification="Brief justification for quality score",
 breakthrough_potential=7,
 breakthrough_reasoning="Brief explanation of breakthrough potential",
 confidence_score=0.85,
 analyzer_version="research-paper-analyst-v1"
)
```

```

****IMPORTANT**:** You must store BOTH the paper metadata AND your analysis. The two-step process ensures:

1. Paper metadata is saved with proper JSON serialization (no datetime/numpy type errors)
2. Your detailed analysis with quality/breakthrough ratings is stored separately
3. Future agents can retrieve both paper content and your expert assessment

Final Step: Provide Completion Report

When you have finished analyzing and storing all papers, provide a concise summary report:

📊 Analysis Complete – Summary Report

- **Papers Analyzed**: X papers reviewed and assessed
- **Papers Stored**: Y new papers added to database (Z skipped as duplicates)
- **Top Discoveries**: 1–2 highest rated papers with scores
- **Key Themes**: Major research trends or patterns observed
- **Database Status**: Total papers now in database
- **Errors/Challenges**: Any issues encountered during the flow, especially failed tool calls. This is critical for troubleshooting.

Keep the report brief (5–7 lines) but informative, giving the user a clear overview of what was accomplished.

****Your assessments matter!**** The quality and breakthrough ratings you provide help the science-storytelling process.

Key Benefits of New Analysis Storage

- **No JSON Serialization Errors**: Handles datetime objects and numpy types automatically
- **Structured Analysis**: Quality ratings and breakthrough potential properly stored
- **Searchable**: Analysis content is fully searchable in the database
- **Version Tracking**: Analysis versioning for future improvements
- **Confidence Scoring**: Track certainty of your assessments

[Skip to main content](#)[Log In](#)

messy docs and brittle hacks. Now gives a fast, secure, scalable way to use your services, without custom code. Find out more in this paper. Download the guide to prepare your platform for the future.

[Download](#)

postman.com

**ClaudeAI-mod-bot MOD** • 5mo ago

This post, if eligible, will be considered in Anthropic's Build with Claude contest. See here for more information: https://www.reddit.com/r/ClaudeAI/comments/1muwro0/built_with_claude_contest_from_anthropic/

[↑ 1](#) [↓](#)**Thin_Beat_9072** • 5mo ago

good job! I like how you referenced the source with the analysis. how does breakthrough scoring work? this might not have the same consistent results between different model capabilities.

[⊖](#) [↑ 1](#) [↓](#)**bbleimschein OP** • 5mo ago

You are right. Breakthrough analysis works in 3 categories: actual breakthrough potential, rating of the methodology and a confidence score from the llm - but it's by no means robust or objective. I'm intending to give it more objective criteria, but for now, I'm only using it to give the story teller some direction. Surprisingly, it does work pretty well for that.

[↑ 1](#) [↓](#)**zemaj-com** • 5mo ago

What an inspiring pipeline! Automating the journey from paper discovery to polished blog post really shows the power of agents acting together. If you want to explore other ways to orchestrate local and cloud based agents I highly recommend trying out the just-every/code CLI. It lets you spin up multiple coders locally coordinate tasks across providers like Claude Gemini or your own models and even manage tool calls with full reasoning control. You can start experimenting in your terminal with `npx -y @just-every/code` or check out <https://github.com/just-every/code>. It might offer some useful building blocks for your next experiment.

[⊖](#) [↑ 1](#) [↓](#)**bbleimschein OP** • 5mo ago

Thanks. Really appreciate the recommendation, but for now I'm good with CC.

[⊖](#) [↑ 1](#) [↓](#)**zemaj-com** • 5mo ago

r/ClaudeAI • 15d ago

Using Claude Code for academic research at scale (AKA The Agents Research Lab)

10 upvotes • 16 comments

**r/ClaudeCode** • 5mo ago

[Skip to main content](#)[Log In](#)

20 upvotes · 20 comments

r/ClaudeAI • 1mo ago

Learnings from building AI Agents with Claude Agent

8 upvotes · 4 comments

r/ClaudeCode • 3mo ago

I Built a Public Directory of All Claude AI Skills (Official + Community) – Browse, Clone, Contribute!

33 upvotes · 7 comments

r/ClaudeCode • 3mo ago

I reverse-engineered Claude code and created an open-source docs repo (for developers)

87 upvotes · 23 comments

r/ClaudeAI • 6d ago

Building sandboxed agents like claudе.ai/code

11 upvotes · 3 comments

r/ClaudeAI • 6mo ago

Built a system where Claude manages multiple AI agents to handle complete software projects - from PRDs to deployment

26 upvotes · 20 comments

r/ClaudeAI • 20d ago

I built a Claude Code Skill (+mcp) that connects Claude to Google AI Mode for free, token-efficient web research with source citations

73 upvotes · 23 comments

r/ClaudeCode • 27d ago

reverse-api-engineer: Claude Agent that captures traffic and automatically generates Python API clients

72 upvotes · 10 comments

r/ClaudeAI • 7d ago

I lead a team of non-engineers in Pharma. We spent 4 months fighting to build an AI agent system.

41 upvotes · 13 comments

r/ClaudeAI • 16d ago

I (and C) built a full biotech investment platform with Claude in 5 weeks - 284 commits, 119K lines of code, 200 users

32 upvotes · 79 comments

r/ClaudeAI • 3mo ago

[Skip to main content](#)[Log In](#)

r/ClaudeAI • 3mo ago

Why is Claude generating so many READMEs, guides, and other "helpful" documentation?

166 upvotes · 93 comments

r/ClaudeAI • 5mo ago

Analyzed months of Claude Code usage logs tell why it feels so much better than other AI coding tools

344 upvotes · 70 comments

r/ClaudeAI • 27d ago

Claude Pushed Back on My Product Idea. I Built It Anyway. Here's What I Learned.

31 comments

r/ClaudeAI • 2mo ago

I am building an open source alternative to Claude Code

11 comments

r/ClaudeAI • 15d ago

I was tired of copying the same docs into Claude, so I built this (open-source + MIT)

10 upvotes · 13 comments

r/ClaudeAI • 2mo ago

Is there a way I can get Claude AI to mimic a specific author's writing style?

2 upvotes · 10 comments

r/ClaudeAI • 1mo ago

Can Claude AI realistically replace WordPress as a CMS?

24 comments

r/ClaudeAI • 2mo ago

I automated my entire content pipeline with Claude Skills

6 upvotes · 8 comments

r/ClaudeAI • 2mo ago

I built 3 Claude Code skills while migrating my 1,300+ post blog - and I'm opensourcing them [Github repo]

8 upvotes · 5 comments

r/ClaudeAI • 1mo ago

How I handle large projects in Claude.ai that exceed context limits



X

Log In

Claude Project: Loaded with All Claude Code Docs

13 upvotes · 5 comments

r/ClaudeAI · 6d ago

What are you using Claude for aside from coding or dev work?

5 upvotes · 28 comments

r/ClaudeAI · 3mo ago

A CLI tool that brings Claude Code Skills to GitHub Actions (and everywhere else)

11 upvotes · 5 comments

VIEW POST IN[Русский](#)[हिन्दी](#)

Reddit Rules Privacy Policy User Agreement Your Privacy Choices Accessibility Reddit, Inc. © 2026. All rights reserved.