

UNIVERSITY OF MANNHEIM
Business School

Research Independence and Quality in the Age of AI

LLMs, AI agents, and what a research data infrastructure has to become

Prof. Dr. Marcel Olbert · University of Mannheim & COBRA

BERD@NFDI Leitungsklausur, ExpLAB · 4 November 2026

Outline

Personal Background

Where I think this is going

Thoughts for the Discussion

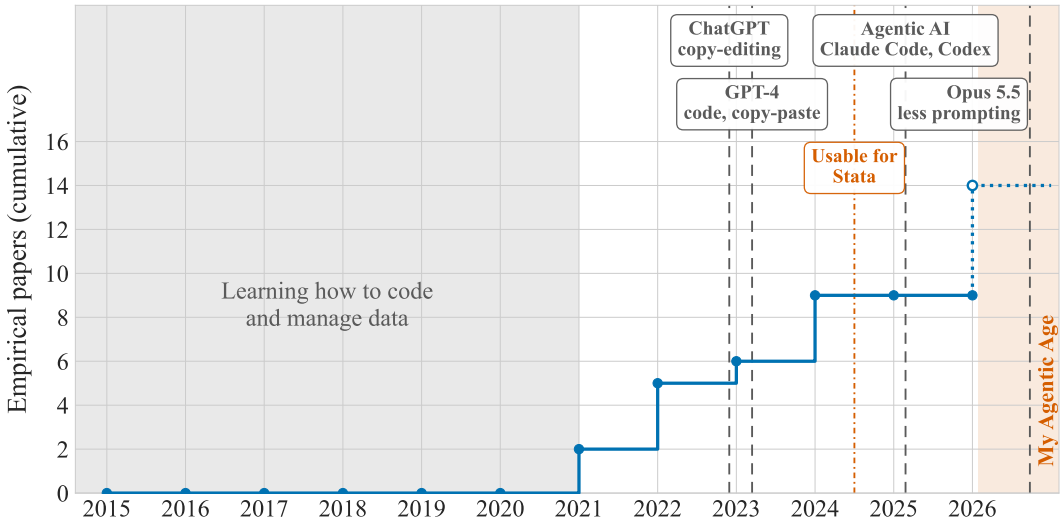
Takeaways

Where I speak from

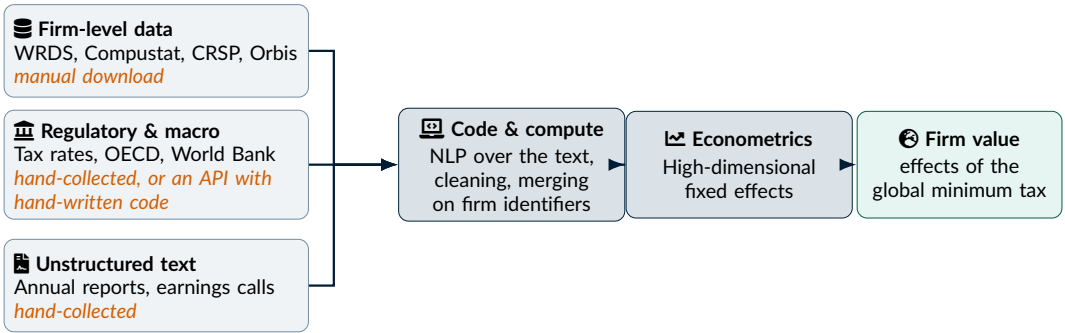
- **Professor of** Taxation, accounting and finance.
- **Empirical researcher**, not an AI engineer. Using firm-level panel data, regulatory parameters, firm disclosures (text and numbers).
- **Academic Director at COBRA**; also establishes a infrastructure: Orbis, textual firm reports, earnings calls, a high-frequency news-based policy tracker.
- **AI Agents in my workflow since 1/2026**: papers, revisions, teaching, admin.

The bar I apply: does it make the research better?

My track record - the future will be different!



The typical empirical paper before agentic AI



My own example: *Measuring the Expected Effects of the Global Tax Reform*, with Roberto Gomez-Cram, *Review of Financial Studies* 2023.

What limited that paper, and what did not

Bottlenecks

- **Coding skills**: NLP, infrastructure, data management
- **Compute**: high-dimensional fixed effects, raw data around 100 GB
- **Access** to costly commercial databases

Sovereignty: not an issue

- A **US setting** needs **US databases** anyway, so the only question was whether to subscribe to a US vendor
- **Hosting, data and code stayed local**: nothing left the building

My setup today: environment and memory

Environment

- **VS Code**: one home for code, LaTeX, and the agent
- **Claude Code**: agent in the terminal; reads, writes, runs
- **Shared Claude Code**, including context layers, on team-workspaces
- **Voice to text** for fast prompting while thinking out loud

Context layers

- Global `CLAUDE.md`: who I am, conventions across projects
- Project `CLAUDE.md`: structure, glossary, rules per project
- **Memories**: persistent, file-based; the agent learns over time
- **Skills**: reusable specialised behaviours, e.g. paper editing

Setup compounds: every new project starts with what the agent already knows about me.

The new top-tier paper after agentic AI

Shocks to the contribution frontier

- The **heavy lifting becomes a commodity**: public unstructured firm-level data, with coding agents collecting, cleaning and verifying it
- **Standard commercial databases** (WRDS included) and public data are **prone to one-shot agentic paper writing**. That is no contribution at a top-tier journal.
- **The competition uses LLMs too**: speed is not an edge

Where the contribution moves

- **Ideas, creative approaches, clever designs**, which is not new
- **More proprietary and administrative data**
- **Measuring the unmeasurable**, pushing the frontier of what can be captured

The new bottlenecks – requirements for an infrastructure

New bottlenecks

- **Cost**: tokens and model subscriptions
- **Data protection and privacy** rule out public models for a growing share of the data

A much stronger case for BERD

- **Professional data management**
- **Local compute capacity**
- **Local LLMs**

Sovereignty issue: **the *third-party* model is now part of the pipeline.**

Three horizons

Short term

Now to 12 months: quick wins on **existing** work, revisions, documentation.

A short window before AI fluency is simply standard.

Medium term

New projects **AI-native from day 1**: clean structure, documented conventions, everything in scripts.

The question becomes: what is my edge, and how does AI amplify it?

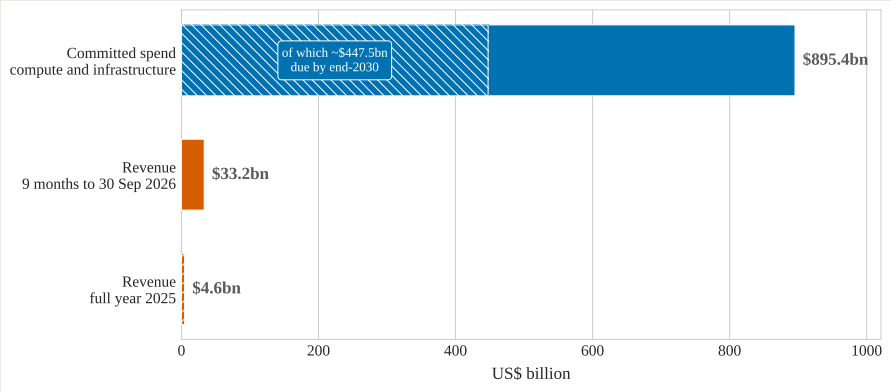
Long term

When everyone has the tool, **ideas and rigour win**.

Non-standard data (proprietary, scraped, hand-collected) becomes *more* valuable, not less.

For infrastructure: **access and provenance outlive any model generation.**

I predict: LLMs and tokens will get more expensive.



Anthropic, Form S-1/A, 2 November 2026; revenue up roughly 1,200% over nine months. As reported by CNBC. Due by end-2030 per Axios.



For 90%+ of our work we use a fraction of what Opus 5.5, or an equivalent model, can do.

- Most research tasks and workflows **standardise**.
- **Reproducibility for journals** forces the same conclusion: the final code has to be **deterministic** anyway. The model helps produce it, or replaces a straightforward manual task.
- Like the average office worker, who uses **under 5% of Excel's functions** and still uses Excel all day.



Sovereignty is an issue now: the key providers are US firms.

- **Geopolitical risk.** Access for non-US researchers can be switched off, throttled or repriced by a decision taken in Washington.
- **GDPR and institutional data protection.** Personal and licensed data cannot simply be sent to a US-hosted model.



Case for local models and local compute: Sovereignty (and cost management).

- But both are **expensive and technically heavy** for an individual researcher or a small team.
- Which is precisely the gap an **infrastructure** exists to fill.

BERD.



Big data is becoming a commodity. Verified, trusted data is not.

- **AI slop**: generated and unverified data, at scale.
- Preventing slop is **infrastructure work**: provenance, verification, curation.
- **Data repositories become more valuable, not less.**



And researchers need new skills. That part is trainable.

- **AI readiness across models**, not fluency in one vendor's tool.
- **Local, deterministic practice**: system prompting, fixed workflows, memories, written conventions. The output has to be defensible, not impressive.
- **This is Academy work. Build on existing formats.**

Takeaways: what the infrastructure should offer

- **Professional data management.** Curated data that carries its provenance. Repositories gain value as analysis is commoditised.
- **Verification, against AI slop.** The quality layer no vendor supplies: provenance, checks, curation.
- **Local compute capacity.** Too expensive and too technical for one chair or a small team.
- **Local models.** A sovereignty answer to US dependence, GDPR and licensed data.
- **AI-readiness training.** Deterministic workflows across models, not tool demos.

Questions for discussion

- Where should a **LLmodel-derived variable** live: in the data portal, in the code repository, or in a record of its own?
- What is the **minimum** a deposit must carry for a 2026 extraction to be checkable in 2031?
- Is **negotiating model processing rights** with data providers something a consortium can realistically do?
- What should BERD **not** do, because the vendors will do it faster?

Thank you

Marcel Olbert · University of Mannheim & COBRA

COBRA

Mannheim Center for Corporate Behavior and Regulation Analysis

Rigorous, independent research on **how businesses respond to economic policy and regulation**.
Seed funding, pre-doc research professionals & visiting programs, data infrastructure, public outreach, conferences and events.

marcel.olbert@uni-mannheim.de

Marcel Olbert on LinkedIn

http://mannheim-business-school.com/cobra