DSTS / DSBS JOINT MEETING · NOVO NORDISK

# $p < 0.05$, Trust Me — *Says the Machine*

Niels Richard Hansen

Department of Mathematical Sciences, University of Copenhagen

19–20 November 2026



Did the statistics profession just die?

The answer may be uncomfortable.

PART 1

What happened?

Timeline

NOV 2022	ChatGPT
MID 2024	Coding agents: tasks of a few minutes[1]
JUL 2025	IMO gold-medal level (OpenAI, Google DeepMind)
FEB 2026	Coding agents: tasks of about 14.5 hours[1]
JUL 2026	IMO: AI systems report perfect scores
JUL 2026	Hugging Face breached by an autonomous AI agent
AUG 2026	OpenAI: ten long-standing open problems, in Lean
SEP 2026	Navier–Stokes blow-up claimed, with Lean proof
OCT 2026	OpenAI releases 722 AI-written math manuscripts

UNDER THE HOOD

2022–23
Pretraining at scale, human feedback

2024
Reasoning models: reinforcement learning on chains of thought

2025
Rewards that can be checked: unit tests, proof checkers

2025–26
Agents: tools, long tasks, many agents in parallel

Navier–Stokes: about 10,000 agents for 88 hours

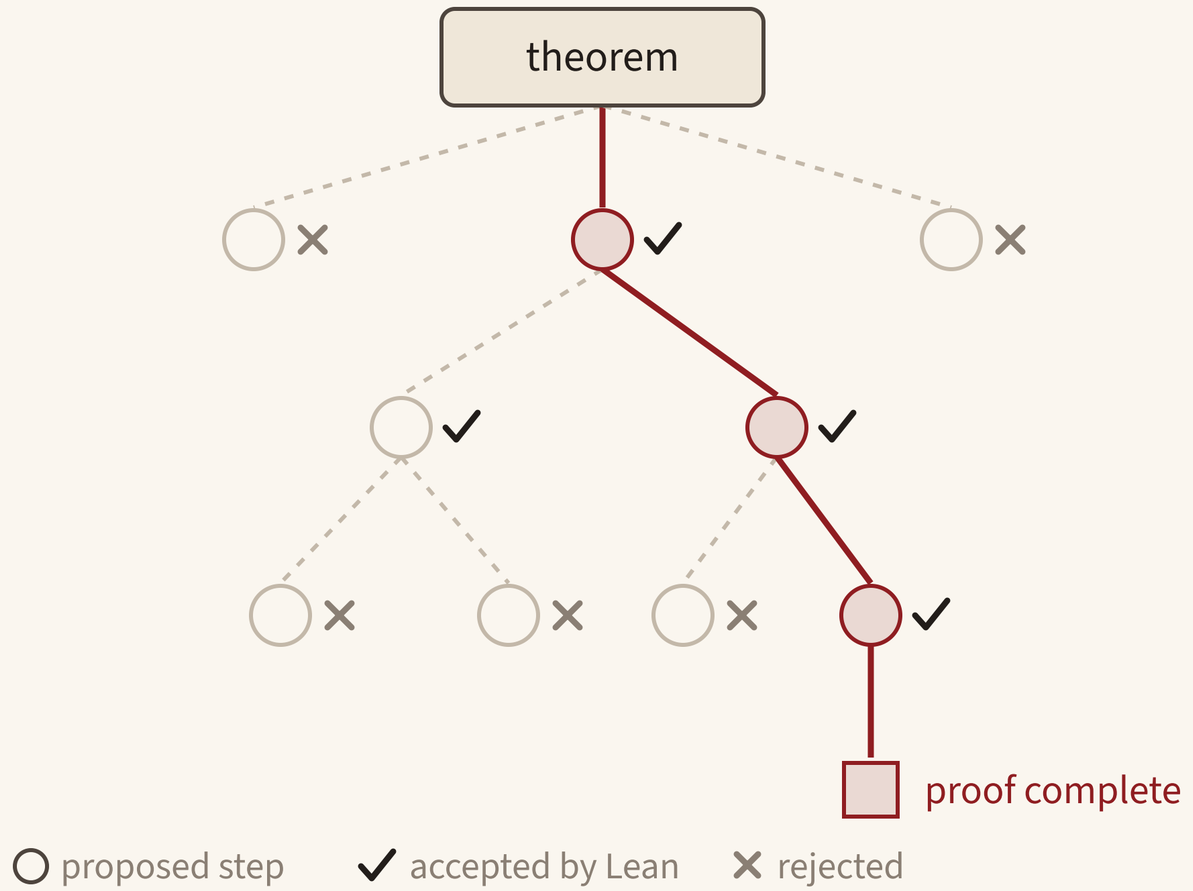
[1] METR time horizon: length of tasks, in skilled-human time, that agents complete half the time. Doubling roughly every four months.

Why math

Professional mathematics is a perfect combination of

- fully digital existence
- correctness is objectively verifiable
- and a large body of high-quality training material

With ***Lean***, mathematical results are fully formalizable and proofs are machine verifiable.



The capability question

WHAT IS SETTLED

- AI systems can be capable of solving highly complex problems
- e.g. doing research, possibly with superhuman capabilities
- and sometimes with scary and dangerous side effects (Hugging Face incident etc.)

WHAT IS UNSETTLED

- The actual business model
- If the resource tradeoff is acceptable
- What we can achieve with small and local models
- Whether the development generalizes to all human capabilities

17 November: a new study shows that coding agents report $p < 0.05$ in between 54% and 98% of simulated null examples.

arXiv:2611.23174

Is the statistics profession affected?

SUBSTANTIALLY ON

- mathematical derivations and proofs
- coding for data analysis, visualization, reporting
- design of studies and methodology selection

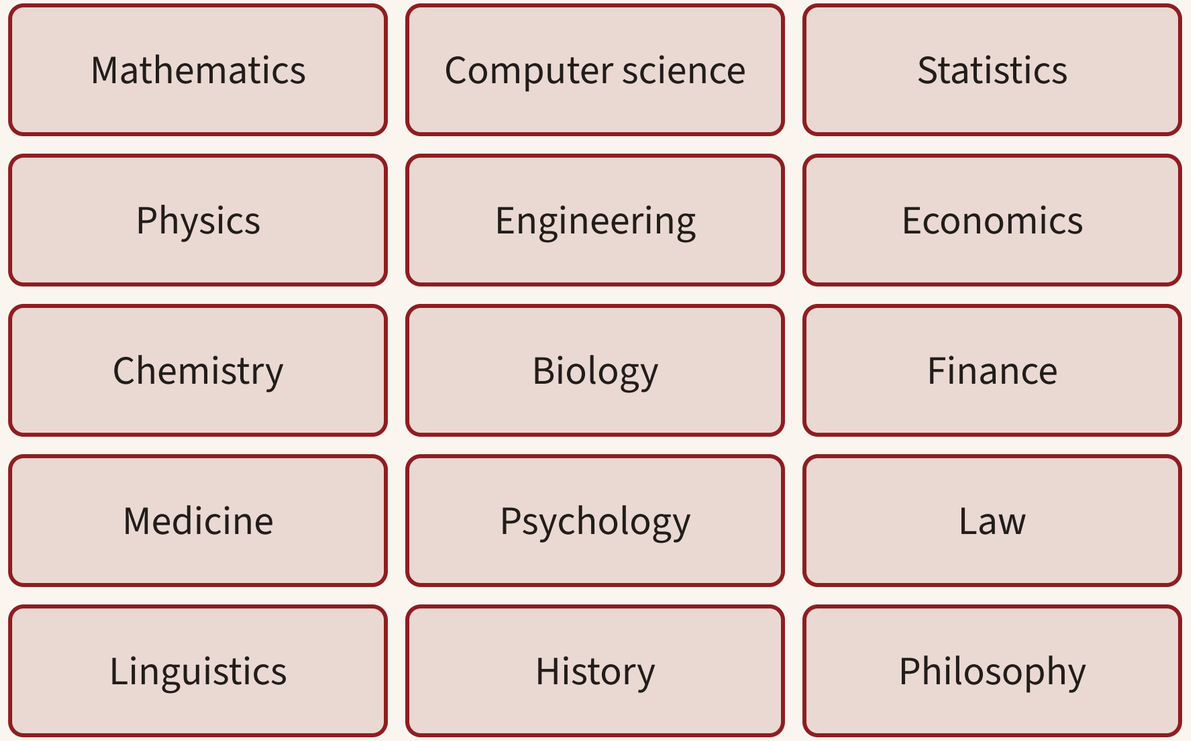
AT THE MOMENT LESS ON

- which questions are worth asking
- who is ultimately accountable
- oversight / regulation

What is next?

The developments will

- spread across all disciplines
- decrease the value of replaceable production skills
- increase productivity (hopefully in terms of quality rather than quantity)
- with dramatic effects on the job market and education system



PART 2

Reactions and positions

Job market reactions

US: junior developers and CS educations

−20% employment of software developers aged 22–25, late 2022 to July 2025

−11% ages 22–25 in the most AI-exposed jobs, Nov 2022 to June 2026 (least exposed: +10%)

−8% undergraduate enrollment in computer science, spring 2026 vs. 2025

Brynjolfsson, Chandar & Chen, Stanford Digital Economy Lab (ADP payroll data; descriptive, not causal). National Student Clearinghouse.

Denmark: firms that adopt AI

≈ 11% below trend by mid-2025: employment in firms that began using AI in 2023/24, vs. comparable firms

Mainly through fewer new hires, concentrated among young and long-educated employees.

Danmarks Nationalbank, analysis, 16 September 2026.

Existential crisis

Programmers, whose identity is writing excellent code

Mathematicians, whose identity is solving difficult problems

Statisticians, whose identity is producing excellent analyses

Production capabilities are greatly devalued when tools can take over.

Is there a version where human value is not attached to productivity?

Some positions among mathematicians

THE SKEPTICAL

- AI is ruining mathematics (leave our field alone)
- Rejection of AI on individual grounds
- The Association for Human Mathematics

AHM, October 2026: 1,008 members; 654 of them pledge to do research without AI.

THE CAUTIOUS BUT OPEN MINDED

- Goals of AI companies are misaligned with those of scientific communities
- AI is a new game-changing instrument: the particle accelerator of mathematics
- Terence Tao: Math 2.0

Tao, Caltech lecture, October 2026: from a scarcity of proofs to an abundance of proofs.

The fragile position

Humans are needed because AI cannot

- write good papers
- do “truly” creative thinking
- have human judgment
- avoid post-selection inference fallacies
- X

Until recently, few mathematicians believed that AI could solve open research problems.

The central question

Do we want a world where no human, not even in principle, can follow and verify a claim or decision made by AI?

PART 3

The defendable position

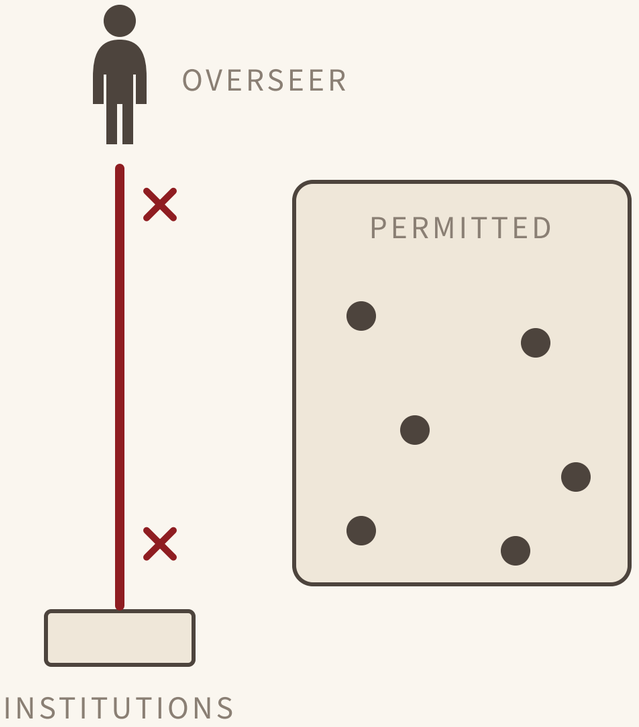
Two roles that we should claim

- ***Regulation*** (what can AI do?)
- ***Oversight*** (does AI comply?)

Implementation of regulation and oversight is a hot topic; it is unsettled if it is even possible with a superhuman AI.

Requires institutional support and the overseer's possibility to reject.

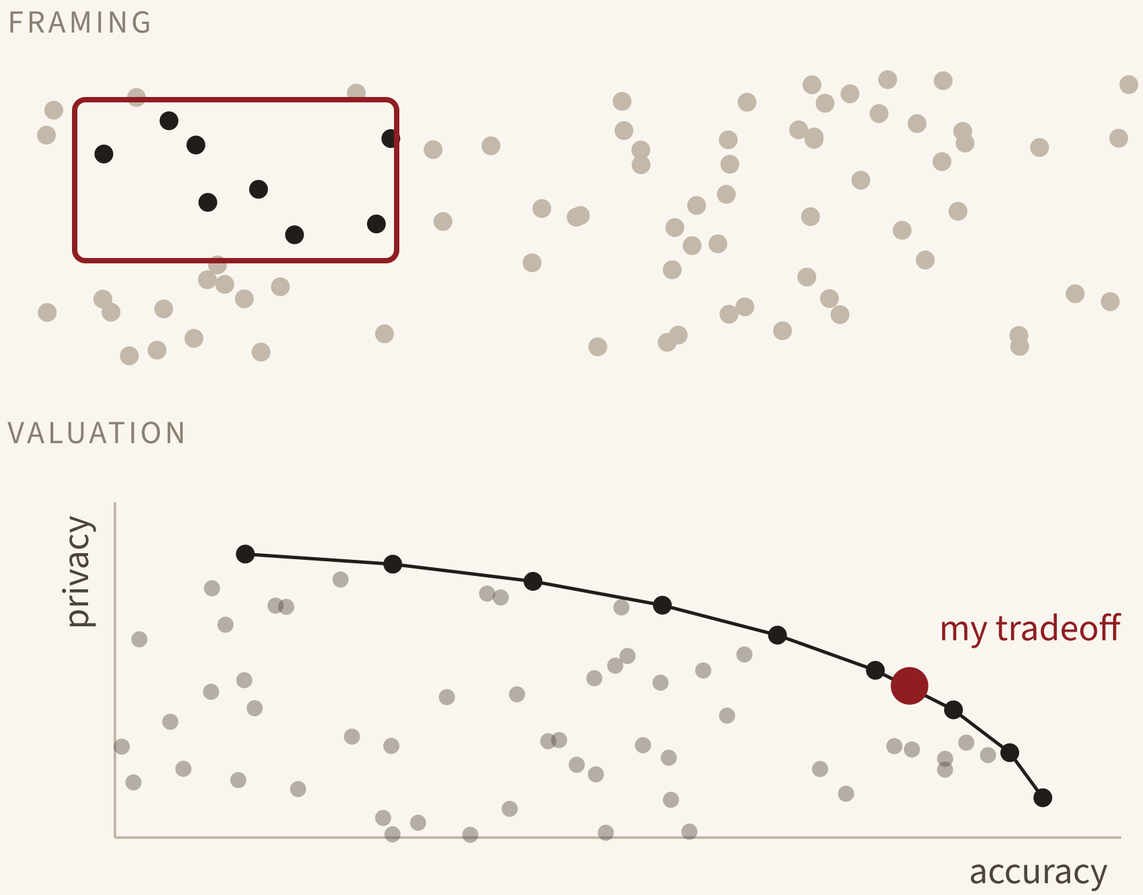




Two other roles—if we want them

We (humans) can insist on the following roles:

- ***Framing***: what is the problem we want to solve
- ***Valuation***: tradeoffs between values with no fixed exchange rate



My position—Humans should maintain agency

- We should demand transparency and decision making that can be followed, justified, questioned and challenged
- We should have the right to frame what the relevant question is
- We should form our own preferences, opinions and values

The $p < 0.05$ claim must be:

> - justified through a transparent analysis, *all the way*

> - the answer to a question that we actually want to ask

> - a claim that we have the right to contest

We do not roll over and do not accept claims by authority alone, only if they can be transparently supported.

Value is detached from productivity and attached to maintaining agency.

PART 4

Outlook

Research expansion

There is an infinity of unsolved problems!

- Accelerated development (proof abundance)
- Unprecedented interdisciplinary research
- Development of new research fields previously impossible
- Human researchers may take roles as: framers, referees, research journalists, or spectators and surveyors.

RISKS

- End of humanity
- Severe disruptions of societal structures
- Unemployment
- Alienation
- AI vs Human credit disputes
- Broken publication model and assignment of merit
- Market mechanisms alone may not grant humans any roles

Education

To ensure competence in our roles we need:

- Education to continue as an institutional structure
- Humans to maintain the framing role of education

- Less tool training (`PROC TTEST`, `t.test()`)
- More enlightenment (what do H_0 and $p < 0.05$ actually mean)
- Sharper focus on limitations and breakdown points
- Increased focus on adversarial behavior
- Critical assessment over productivity tooling

RISKS

- uncertain job market connection
- individual incentives are credentials, not agency
- uncertain human educator roles
- what we actually want is more difficult to measure