What changes when a mathematician works with an AI agent?

Özge Şahin
Delft Institute of Applied Mathematics, TU Delft
KWG Math & AI day, Utrecht

23.10.2026

Who am I?

TU Munich — MSc, PhD 2023

DIAM, TU Delft — since 2023

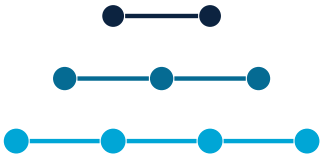
using LLMs since November 2022

applied statistics



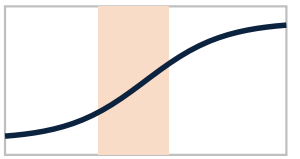
Healthcare and finance

computational statistics



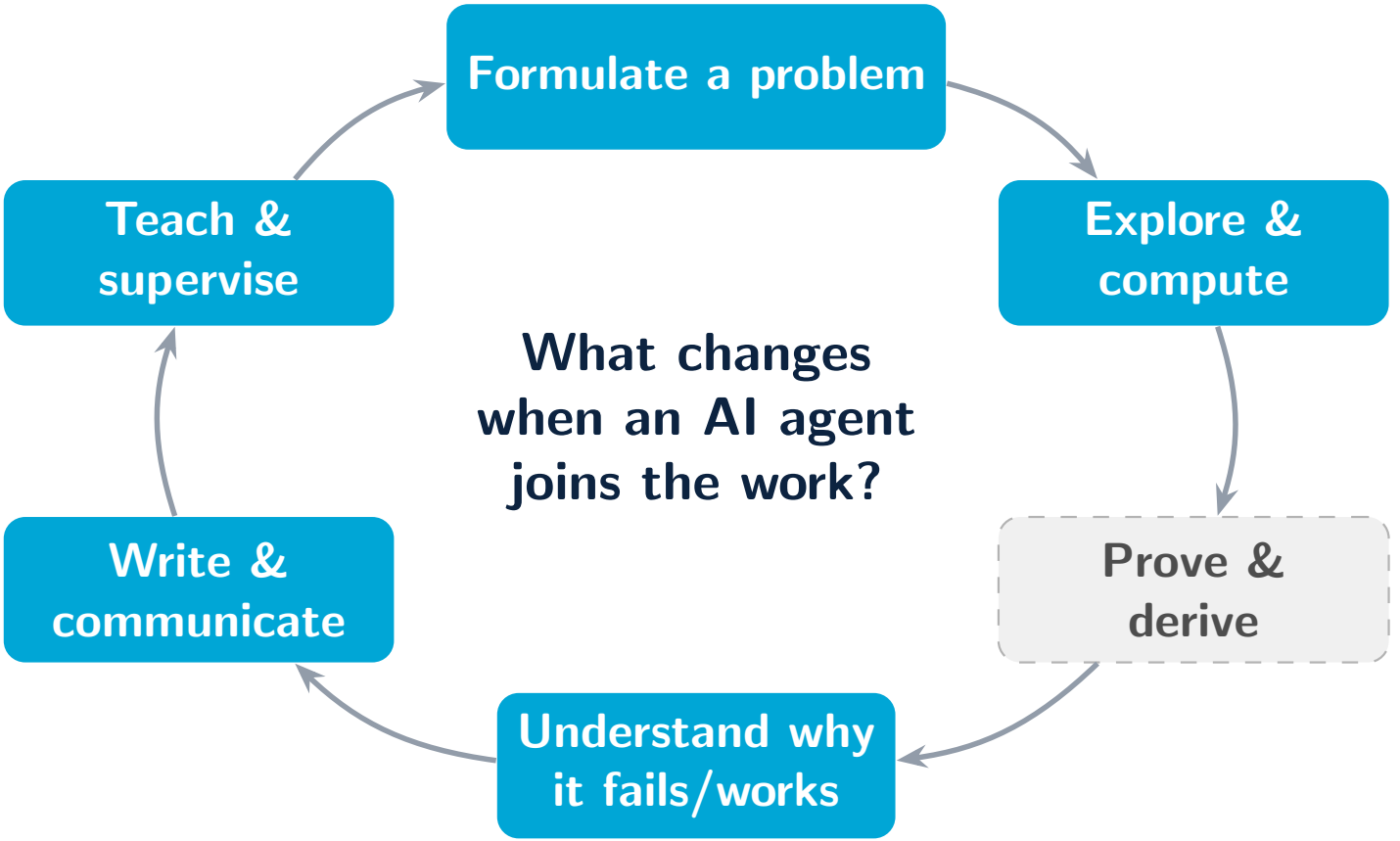
Tractable dependence modeling tools

mathematical statistics



Multivariate Statistics

More than proving theorems



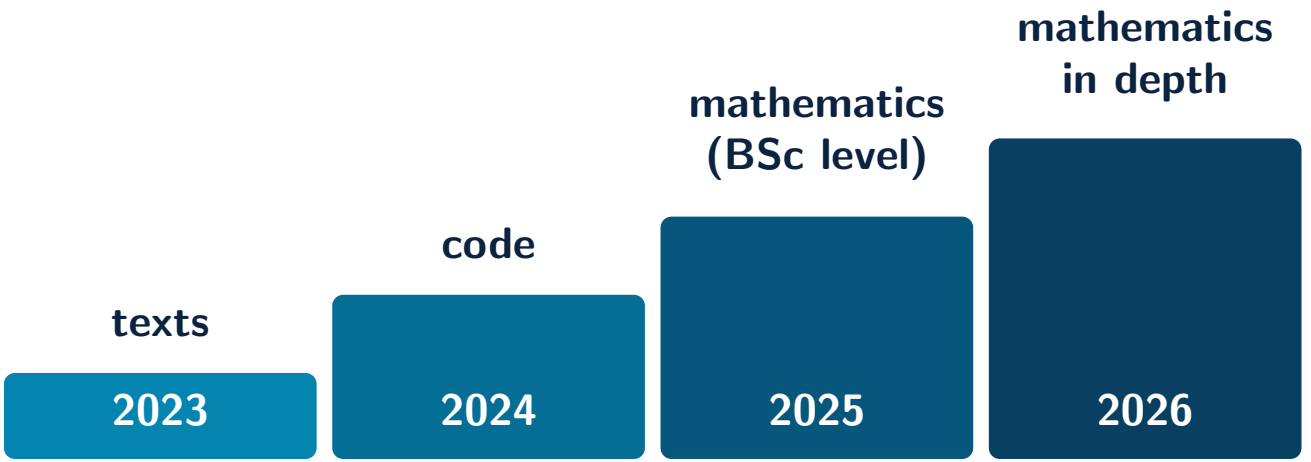
Specific questions

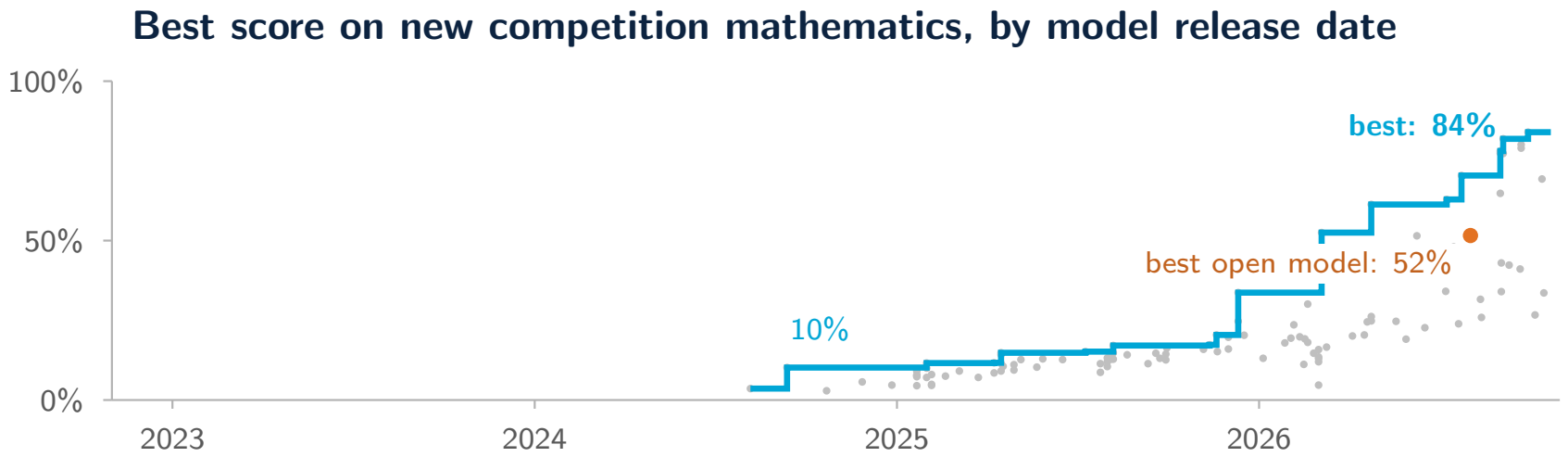


1. How did my way of working change?
2. What is my role?
3. What comes for teaching?
4. How can a department support this?

From November 2022 to October 2026

Q1: How did my way of working change?





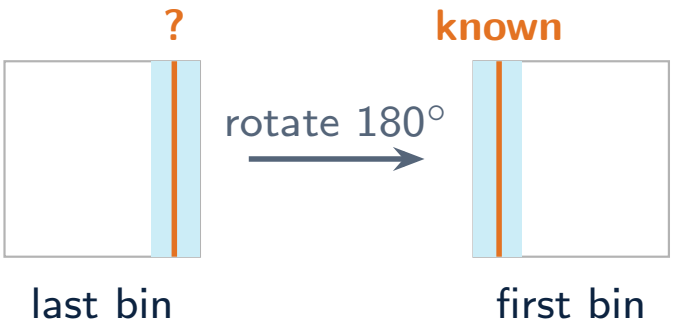
MathArena (ETH Zürich), 112 models, by 22 October 2026; Epoch AI tracks over 3600 AI models now.

A proof idea

Q1: How did my way of working change?

My request (April 2026): the referee's question: if we can do the first bin of Clayton and Frank, can't we also do the last bin?

I supplied my folder to an LLM: two books on copulas, related papers, my manuscript, and the referee report



Its contribution: from my folder, a proof in TeX that I could not think of

My contribution: the background, and checking each step

Connecting dots

Q1: How did my way of working change?

My request (July 2026): explain a pattern in my simulations

I supplied my folder to Codex: my manuscript and its simulation results



one correlation per group

correlation changes $\Rightarrow$ worse fit

worse fit $\not\Rightarrow$ correlation changes

Its contribution: a proof for a pattern in my simulations

My contribution: checking every step

Computational efficiency

Q1: How did my way of working change?

My request (July 2026): this code takes time; can I make it more efficient with minimal touches?



I supplied my folder to Codex: my R code

before: 640 000 density calls per setting

```
for(i in 1:N^2){        # 160 000 nodes
  ctilde_xy = dbicop(u = c(u_i,u_j), ...)
  c_xy      = dbicop(u = c(u_i,u_j), ...)
  ...
}
# whole function called twice
```

after: 2 density calls per setting

```
c_xy      = dbicop(u = quad_grid$u, ...)
ctilde_xy = dbicop(u = quad_grid$u, ...)
ref_0 = compute_direction(c_xy, ctilde_xy)
ref_1 = compute_direction(ctilde_xy, c_xy)
log(true_c) - log(input_c)
sqrt(max(variance, 0))
```

Its contribution: $\approx$ 2 days $\rightarrow \approx$ 20 min, same numbers but less overhead

My contribution: spotting the two changes I did not ask for (orange)

My talks with my own style

Q1: How did my way of working change?

My request (September 2026): make a 40-minute talk from our manuscript, in the style of my three talks ... (a detailed and long prompt)

I supplied my folder to Codex: three of my talks that I liked, and our manuscript

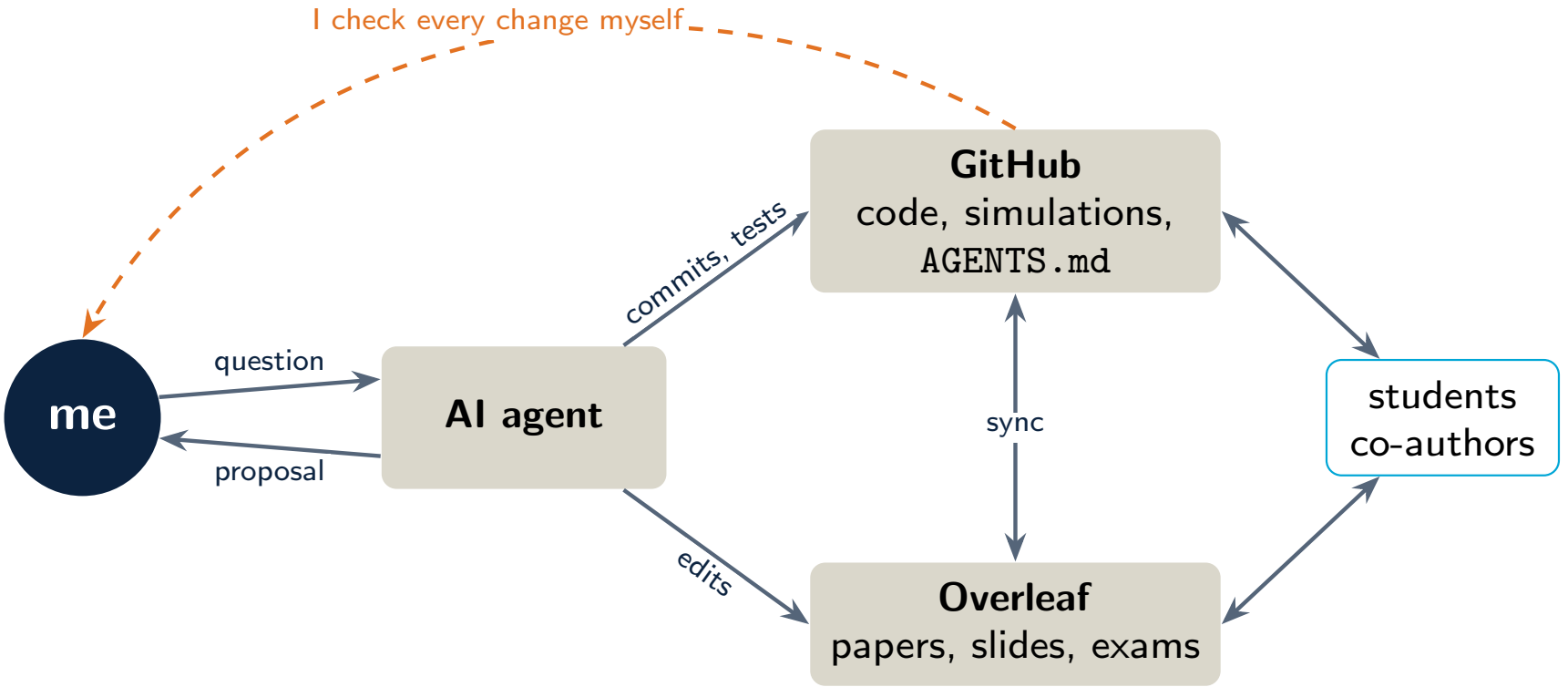


Its contribution: a talk in my style: my opening questions, my colours and theme

My contribution: choosing the examples, and checking every number against the paper

The full process

Q2: What is my role?

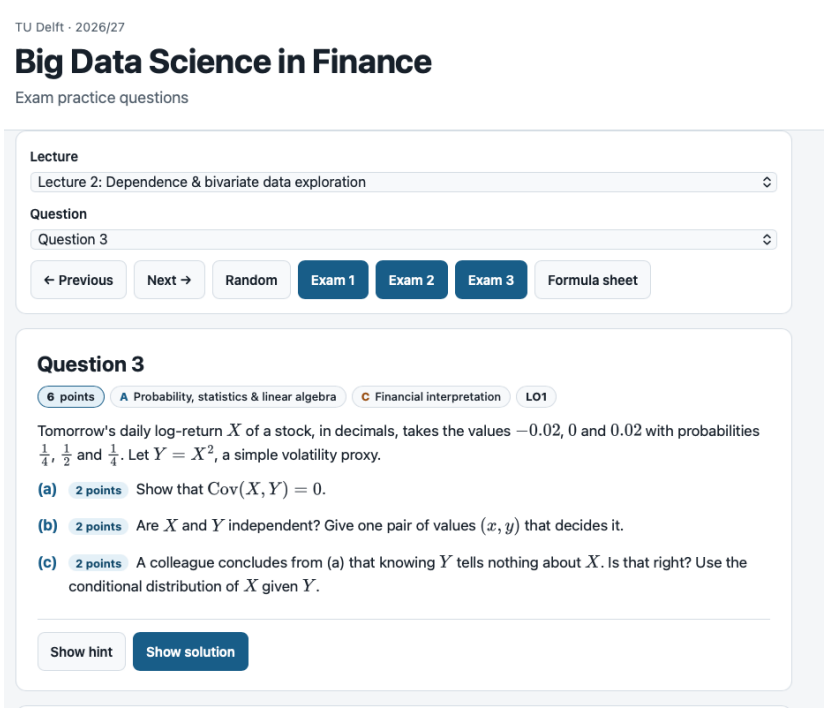


Teaching: a vibe-coded practice exam site

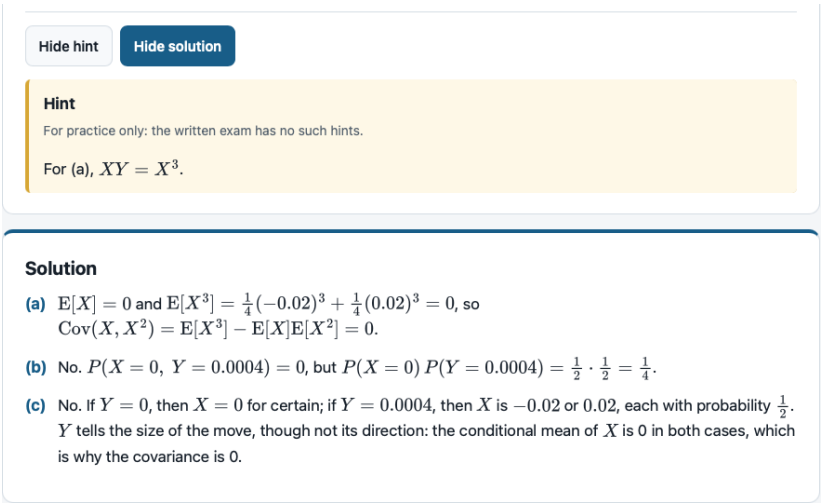
Q3: What comes for teaching?

39 questions **3** mock exams

1. Work on paper



2. Hint and worked solution

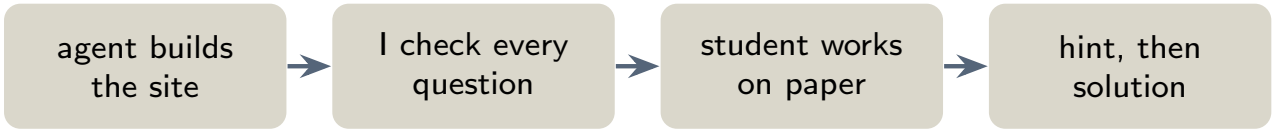


oezgesahin.github.io/dsf_exam_26_27/

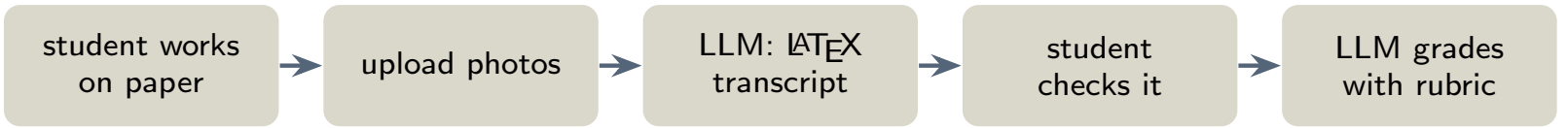
Different designs for student feedback

Q3: What comes for teaching?

A website (my course)



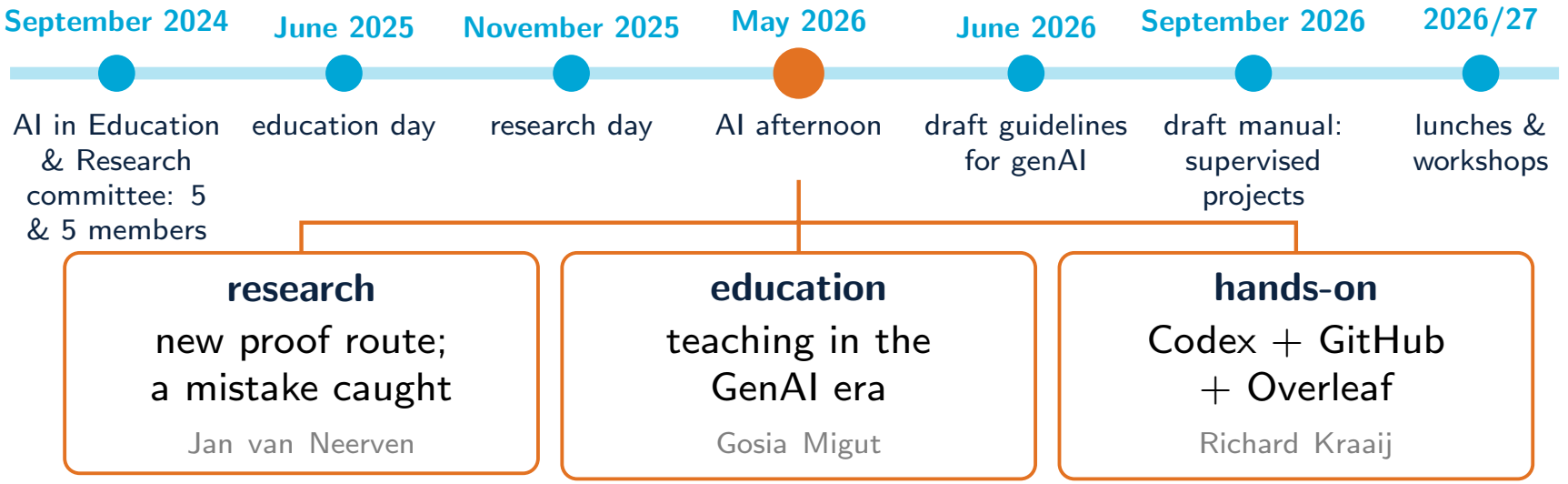
B AI feedback (Real Analysis, Emiel Lorist)



	A	B
cost per answer	**none**	**API cost**
student data	**none**	**sent to a company**
feedback	**generic**	**on own work**

AI at DIAM so far

Q4: How can a department support this?



- Personal AI budget and time?
- Open source models as competition?

Next: applied mathematics

My ideas for the future

ESG ratings
with finance experts

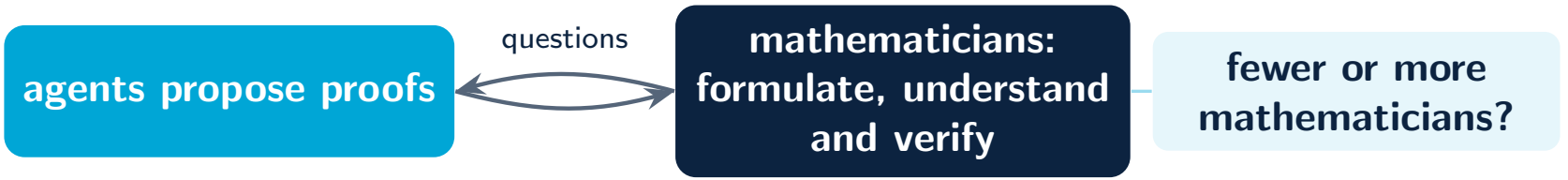


patient risk after bowel surgery
with clinicians

1. The analysis gets faster and more; formulating problems with field experts usually does not.
2. More time to think deeply about the problem and its link to society.

Next: mathematics

My ideas for the future



Who pays: states or companies? I believe we need open source models as competition. If the companies compete with each other, will they still publish their results without being formulated/verified/understood by mathematicians?

How we write mathematics:



My summary

1. It proposes proofs and carries out whole tasks from my examples; **my part is the question and review**.
2. Theoretical and computational work takes less time than before; **applied work needs some time with field experts**.
3. Departments: **time**, a **personal AI budget**, **open source** and public alternatives.

Thank you! (O.Sahin@tudelft.nl)