

Comments on Automation, Learning, and Career Dynamics (Afrouzi, Blanco, Drenik, Hurst)

Jon Kleinberg

Cornell University



Speculating about the AI Future

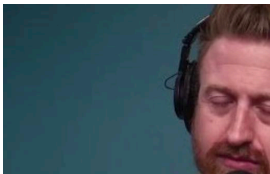
What's next in AI: 7 trends to watch in 2026

"Engineers don't write code anymore" — Anthropic CEO says AI is about to eat the entire profession

85 Predictions for AI and the Law in 2026

News By [Kevin Okemwa](#) published Jan 5, 2026 by: [Oliver Roberts](#), AI & the Law Newsletter

© Posted On Monday, January 5, 2026



RELATED PRACTICES & JURISDICTIONS

Artificial Intelligence
Law Office Management
All Federal
All States

AI's Mainstream Move: What Music Looks Like in 2026

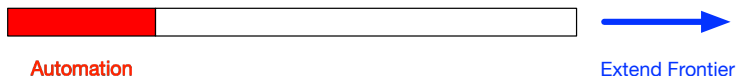


@hardfork

Subscribe

Will A.I. Finally Crack a Millennium Prize Problem in 2026?

The current paper



Automation, Learning, and Career Dynamics
(Afrouzi, Blanco, Drenik, Hurst)

Two kinds of agents:

- Workers perform low-level tasks and also learn by doing.
- Managers extend the frontier of tasks.

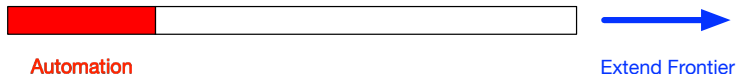
AI (the underlying technology) interacts with both kinds of agents:

- AI automates some set of lower-level tasks done by workers.
- AI can help managers expand the frontier of available tasks.

Two kinds of equilibria:

- Higher levels of automation with lower levels of learning
- Higher levels of learning with lower levels of automation

Some thoughts about extensions



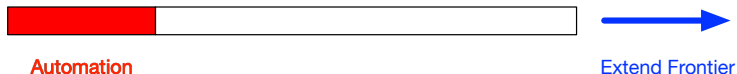
Worker-to-manager career ladder gives the agents a dynamic trajectory.

- Should the AI have any kind of dynamic component?
- In addition to the fact of its rapid change, a technology like this almost must change to be responsive to changes in the world.

Some possible implications of a dynamically changing AI:

- Workers already make two contributions: performing tasks, and learning. Is providing training data for the AI a third contribution they make?
- Managers currently stop learning.
But if the AI is dynamically changing, do they need to learn as it changes?

Some thoughts about extensions



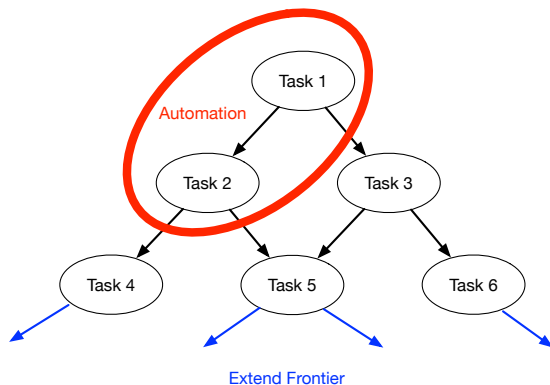
Currently automation is all-or-nothing at the low end:
each task is either done by AI or by a worker.

Compare with a large literature on human-AI collaboration.

[Bansal et al 2019, Chan et al 2019, Lubars and Tan 2019, Lai et al 2022, Donahue-Chouldechova-Kenthapadi 2022, Alur-Raghavan-Shah 2024, De Toni et al 2024, Iakovlev and Liang 2024, Lykouris and Weng 2024, Vafa et al 2024, Xu 2024, Dean-Dong-Jagadeesan-Liu 25, Greenwood et al 25, Noti et al 25]

- Empirical evidence mixed and domain-dependent.
- What could a formalism for worker-AI collaboration do in this model?
- What do workers learn from collaboration with AI?
- If AI sets a floor on outcome quality, does this make worker promotion decisions more difficult?

Dependent Tasks and the Handoff Problem



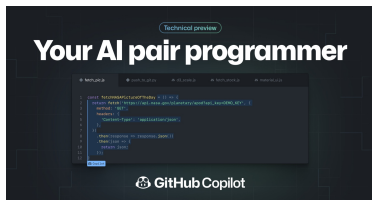
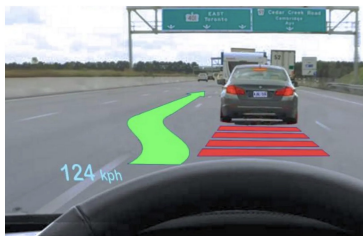
Tasks have dependencies.

- AI and managers don't simply extend a frontier linearly. They have a search problem in a directed graph of dependent tasks.

Once we have dependencies, we have a *handoff problem*.

(Goldenfein et al 2020, Hamade et al 2024)

The Challenge of Handoffs from AI to Humans



Danger with human-AI collaboration: handoff from AI to human.

- AI is in charge of Task i .
- Human is in charge of Task j , which depends on Task i .
- Often the solution to Task i is not sufficient:
The human needs to know how to build on it to complete Task j .
- Related: Goldenfein-Mulligan-Nissenbaum-Ju 2020.

Better design goal: AI is powerful but also produces solutions that the human knows how to build on.

- A kind of interpretability through actions.

Chess as a long-standing model system for decision-making

- “The drosophila of psychology”
—Herb Simon and William Chase, 1973.
- “The drosophila of artificial intelligence”
—John McCarthy, 1990.



A new dimension: Computers have far surpassed the best human chess players.

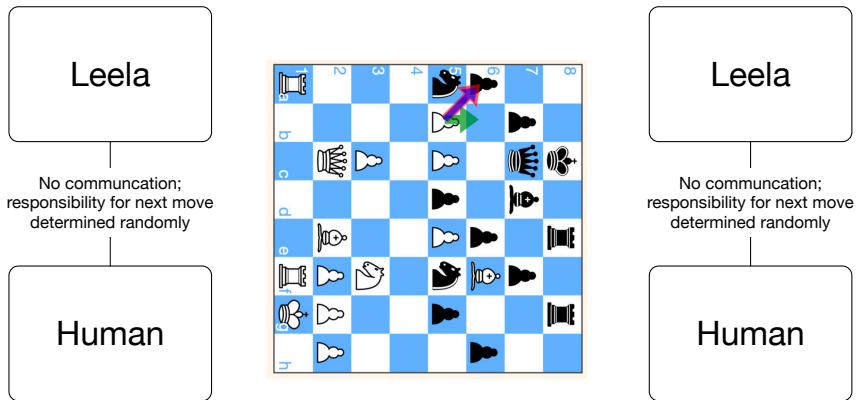
- 1988: First recorded win by computer against human grandmaster under standard tournament conditions.
- 1997: Deep Blue defeats world champion Kasparov in 6-game match.
- 2002-2003: Draws against world champions using desktop computers.
- 2005: Last recorded win by a human player against a full-strength desktop computer engine under standard tournament conditions.
- 2014-17: End of “Centaur” competitions with human-machine teams.
- 2018: AlphaZero trained using self-play; open-source version Leela.

Elo ratings: 1100 = competent human player //

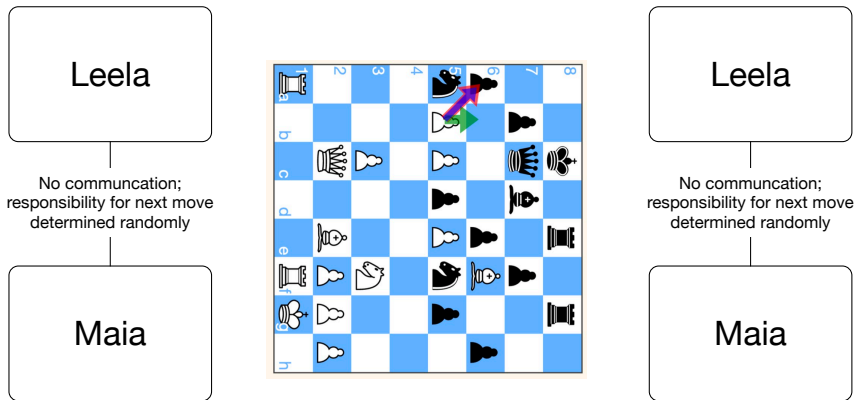
1800 = strong amateur // 2300 = very strong amateur //

2800 = best human players // > 3600 = current chess engines.

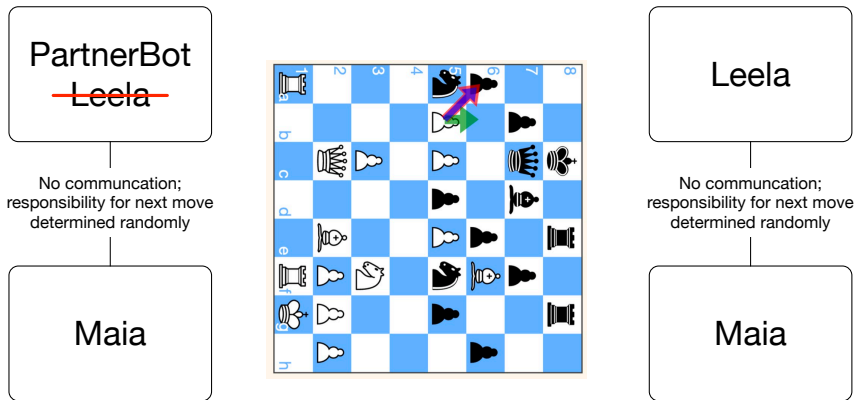
A Handoff Experiment Using Chess



A Handoff Experiment Using Chess

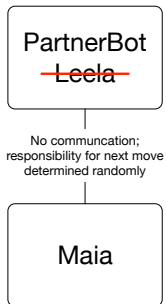
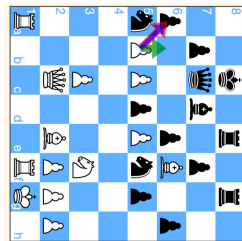
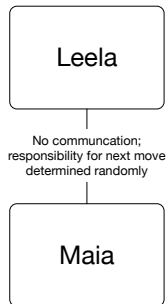


A Handoff Experiment Using Chess



Performance

	Leela	PartnerBot	Maia
Teamed with Maia vs. Leela-Maia	.500	.665	.075
In solo competition vs. Leela	.500	.140	.000

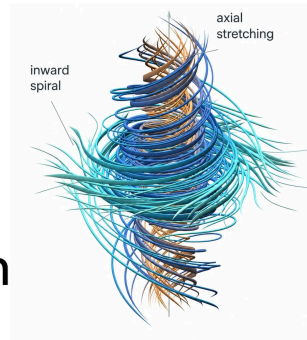


Can We Apply This Perspective to What's Happening in Mathematics?

h Products Business Developers Company Foundation Q

September 8, 2026 Research Publication

On the Navier–Stokes Millennium Prize Problem



Can We Apply This Perspective to What's Happening in Mathematics?

A Severe Misalignment of AI in Mathematics

Published 11 September 2026 · DOI [10.5281/zenodo.22737750](https://doi.org/10.5281/zenodo.22737750) ▶ English

Over the last few months, the mathematical capabilities of LLMs have improved dramatically, to the point that they can solve major outstanding problems in many fields of mathematics. However, the push by AI companies to solve mathematical problems as a benchmark is detrimental to the science of mathematics, and to the mathematical community. The goals of the AI companies and the goals of the mathematical community are severely misaligned. We see these as part of broader alignment issues impacting other scientific and creative professions, as well as the whole of society.

ADD YOUR NAME

If you support this declaration, you are invited to endorse it.

Endorse the Declaration

Your endorsement appears right after you sign in with ORCID, or after you confirm your academic email address.

Can We Apply This Perspective to What's Happening in Mathematics?

The mathematical community functions, in many ways, as a miniature version of humanity. It consists of individuals using a wide variety of different approaches, joined by core values. The most precious resources of our profession are students and ideas, and these we nurture with great care. We feel responsible to let them grow to their full potential, until they can live a life of their own in the mathematical world. For students we often suggest problems with the core intention of developing skills making them well-positioned for advances in research and elsewhere. Our ideas we disseminate in talks, private discussions and careful writeups, connecting them to the previous ideas of others. These processes invariably take time and are based on human interaction.

Can We Apply This Perspective to What's Happening in Mathematics?

Problem-solving versus theory-building

- Analogous to workers versus managers? But not traditionally a ladder.
- W.T. Thurston, *On Proof and Progress in Mathematics*:
*"The question is not "How do mathematicians make progress in mathematics?"
Rather, as a more explicit (and leading) form of the question, I prefer "How do mathematicians advance human understanding of mathematics?"*
- Problem-solving known to be fully automatable since circa 1900, but until now had relatively little practical impact on research in mathematics.
- The feasibility of automated problem-solving has grown abruptly.
- The set of all provable theorems has a dependency (handoff) structure.

Several existential fears:

- The demolition of the training pipeline.
- Handoff from incomprehensible machine proofs.
- The possibility that theory-building could get automated too.
 $\psi = 1$ in the language of the current paper.