This presentation was created for educational purposes.

The Gen AI Adoption Framework

Society for Insurance Financial Management (SIFM)
December 3, 2025

Frank Schmid
Chief Technology Officer

Contents

Co-Innovation and Formation of Organizational Capital

Frank Schmid, "60, 30, 15, ... 7.5? What History Suggests About Generative AI," Gen Re, December 2, 2025, https://www.genre.com/us/knowledge/contributors/frank-schmid?page=1&facet=all.

Frank Schmid, "The Potential of AI for Process Innovation," Gen Re, September 4, 2025, https://www.genre.com/us/knowledge/publications/2025/september/the-potential-of-ai-for-process-innovation-en.

Frank Schmid, "Generative Artificial Intelligence in Insurance – Three Lessons for Transformation from Past Arrivals of General-Purpose Technologies," Gen Re, March 17, 2024, https://www.genre.com/us/knowledge/publications/2024/march/generative-artificial-intelligence-in-insurance-2-en.

The Defining Characteristics of a GPT

"Whole eras of technical progress and economic growth appear to be driven by ... GPTs, [which are] characterized by **pervasiveness** (they are used as inputs by many downstream sectors), **inherent potential for technical improvements**, and '**innovational complementarities**,' meaning that the productivity of R&D in downstream sectors increases as a consequence of innovation in the GPT. Thus, as GPTs improve they spread through the economy, bringing about generalized productivity gains."

--Bresnahan and Trajtenberg, 1992.

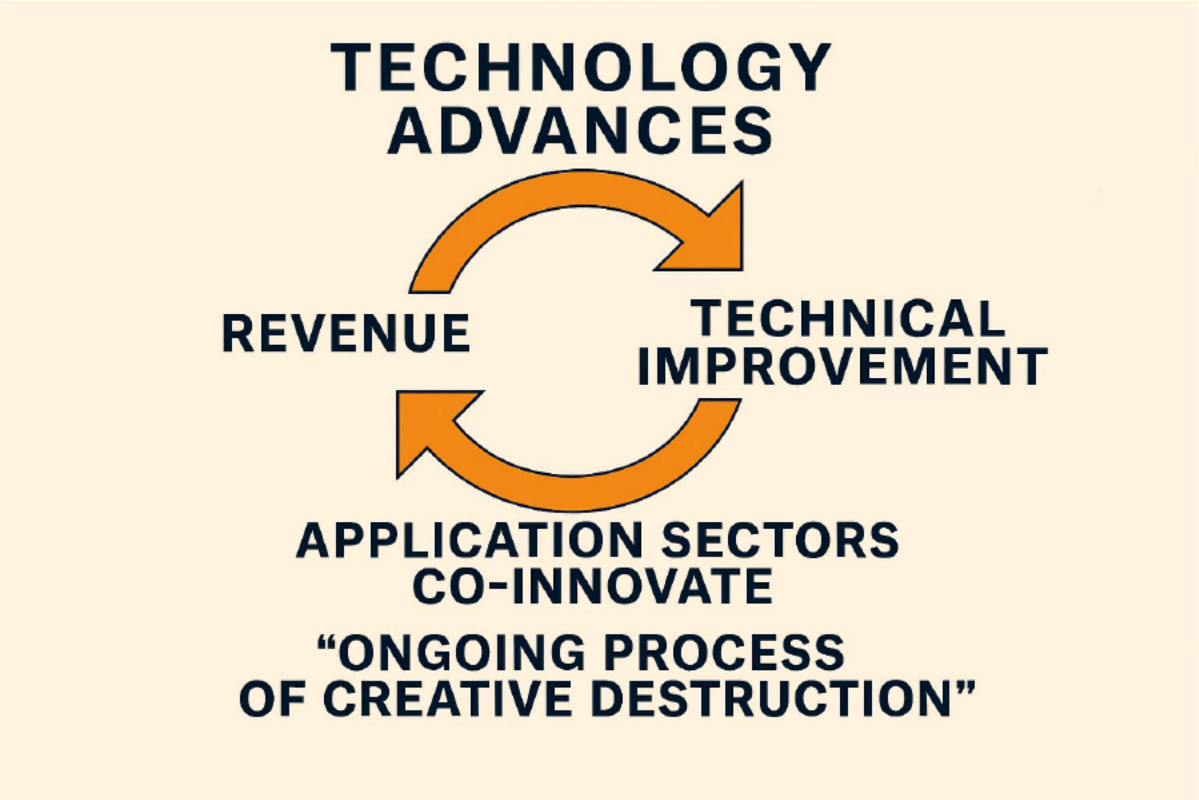
See Timothy F. Bresnahan and Manuel Trajtenberg, "General Purpose Technologies: 'Engines of Growth'?," NBER Working Paper No. 4148, August 1992, https://www.nber.org/system/files/working_papers/w4148/w4148.pdf. Quoted from Erik Brynjolfsson and Andrew McAfee, *Race Against the Machine*: *How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*, Lexington (Mass.): Digital Frontier Press, 2011, Section "Computing the Economy: The Economic Power of General Purpose Technologies."

Bresnahan and Trajtenberg (1992) identify as GPTs the steam engine, the electric motor, and semiconductors & computers.

Proprietary Information – Not for Distribution | © General Reinsurance Corporation

The Dynamics of Gen AI Adoption at the Level of the Economy

Feedback loop of technological advances and co-innovation in the application sectors



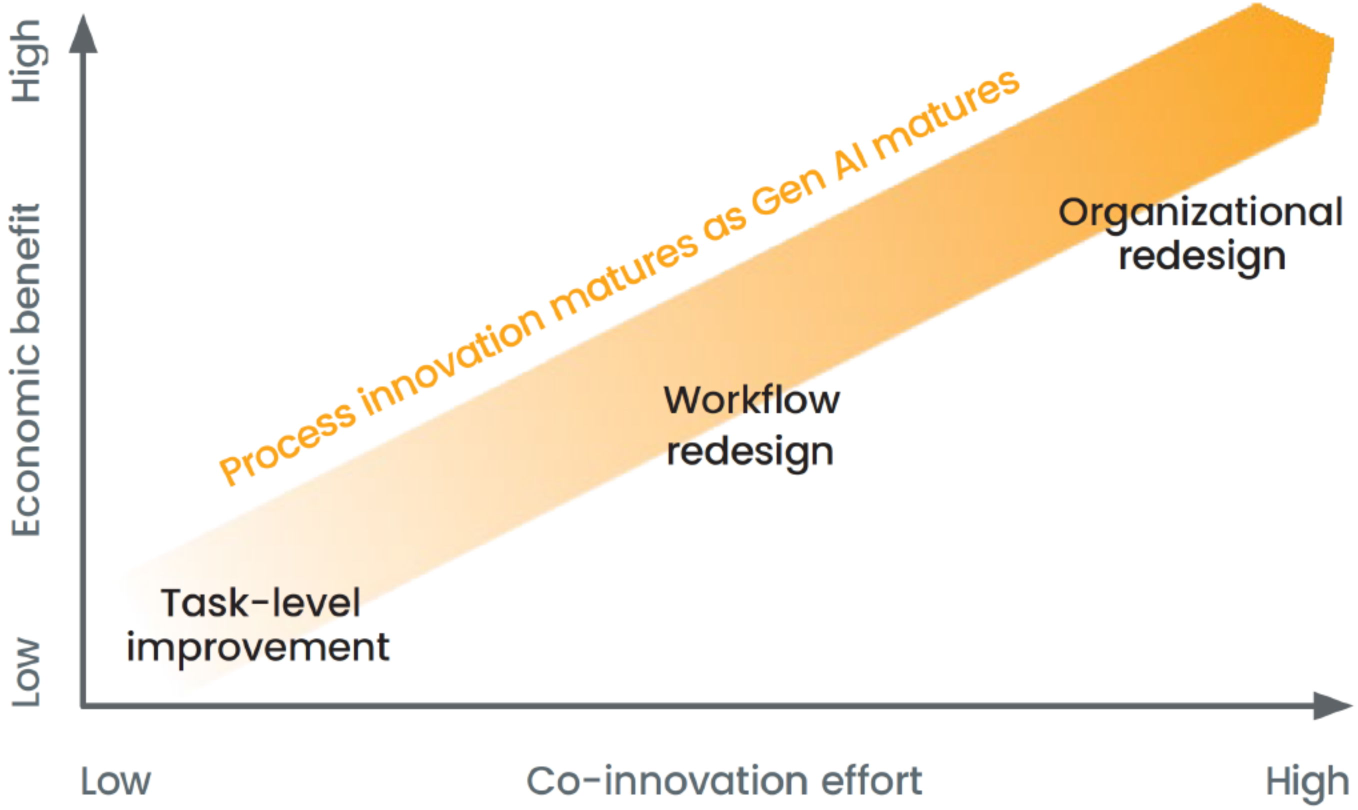
The statement Brynjolfsson and McAfee (2011) made about digitization also applies to the adoption of generative AI, "in other words, it is not a single project providing one-time benefits. Instead, it's an ongoing process of creative destruction; innovators use both new and established technologies to make deep changes at the level of the task, the job, and process, even the organization itself."

Source: Gen Re.

See Erik Brynjolfsson and Andrew McAfee, *Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*, Lexington (Mass.): Digital Frontier Press, 2011, Section "Computing the Economy: The Economic Power of General Purpose Technologies."

The Dynamics of Gen AI Adoption at the Level of the Firm

From task-level improvement to workflow redesign, on to organizational redesign



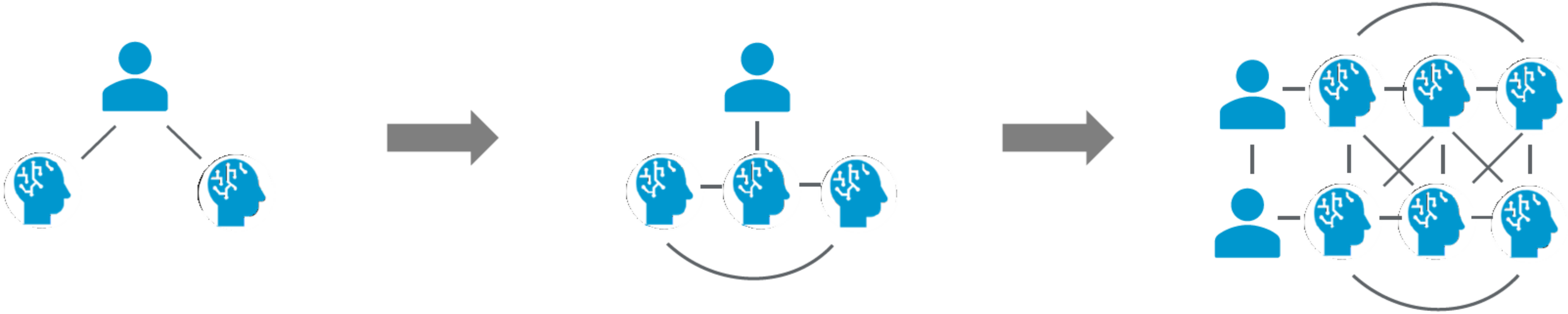
The statement Brynjolfsson and McAfee (2011) made about digitization also applies to the adoption of generative AI, "in other words, it is not a single project providing one-time benefits. Instead, it's an ongoing process of creative destruction; innovators use both new and established technologies to make deep changes at the level of the task, the job, and process, even the organization itself."

Source: Gen Re.

See Erik Brynjolfsson and Andrew McAfee, *Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*, Lexington (Mass.): Digital Frontier Press, 2011, Section "Computing the Economy: The Economic Power of General Purpose Technologies."

The Journey to the Frontier Firm

Humans with Assistant(s) | **Human-Agent Teams** | **Human-Led, Agent-Operated**



An empirical study on computerization demonstrated how technology reshaped job skill demands by replacing routine cognitive and manual tasks while complementing nonroutine problem-solving and complex communication. This shift reduced routine labor, increased nonroutine cognitive work, and drove demand for college-educated labor from 1970–1998. Notably, **more than half of this change in skill demand came from evolving task content within existing occupations**, meaning jobs transformed internally rather than simply disappearing.[1]

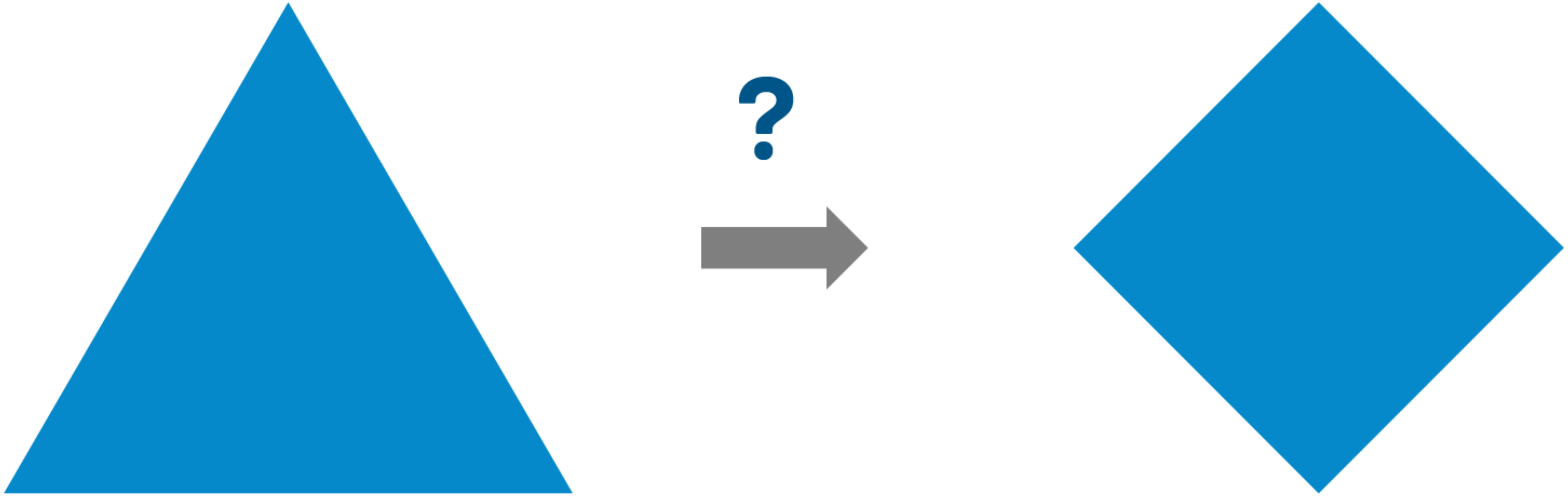
See Microsoft, "2025: The Year the Frontier Firm is Born," *Work Trend Index Annual Report*, April 23, 2025, https://www.microsoft.com/en-us/worklab/work-trend-index/2025-the-year-the-frontier-firm-is-born.

1) David H. Autor, Frank Levy, and Richard J. Murnane, "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, Vol. 118, No. 4 (November 2003), 1279-1333, https://economics.mit.edu/sites/default/files/publications/the%20skill%20content%202003.pdf.

Proprietary Information – Not for Distribution | © General Reinsurance Corporation

From Pyramid to Diamond?

Companies may move from a pyramid shape (large base of entry-level roles) to a diamond shape, with fewer low-skill positions, a strong middle layer of specialized and technical roles, and a continued need for strategic leadership at the top.



See, for instance, Robert Buckland, "AI will disrupt equity research from the bottom up," *Financial Times*, September 20, 2025, https://www.ft.com/content/137ed8ea-5711-4b11-8458-44152fb44990.

On evolving job market implications, see Erik Brynjolfsson, Bharat Chandar, and Ruyu Chen, "Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence," August 25, 2025, https://digitaleconomy.stanford.edu/wp-content/uploads/2025/08/Canaries_BrynjolfssonChandarChen.pdf. See also Justin Lahart, "There Is Now Clearer Evidence AI Is Wrecking Young Americans' Job Prospects," *The Wall Street Journal*, August 26, 2025, https://www.wsj.com/economy/jobs/ai-entry-level-job-impact-5c687c84.

Impact of Co-Innovation on Organizational Capital

Empirical evidence from computerization (1987 – 1997)

Study Overview. Brynjolfsson, Hitt, and Yang (2002) examined 1,216 publicly traded U.S. firms to quantify organizational capital complementing computer assets.

Organizational Capital. Organizational capital is intangible and estimated using the firm's ratio of market value to book value (Tobin's q), reflecting investments in human capital and organizational traits.

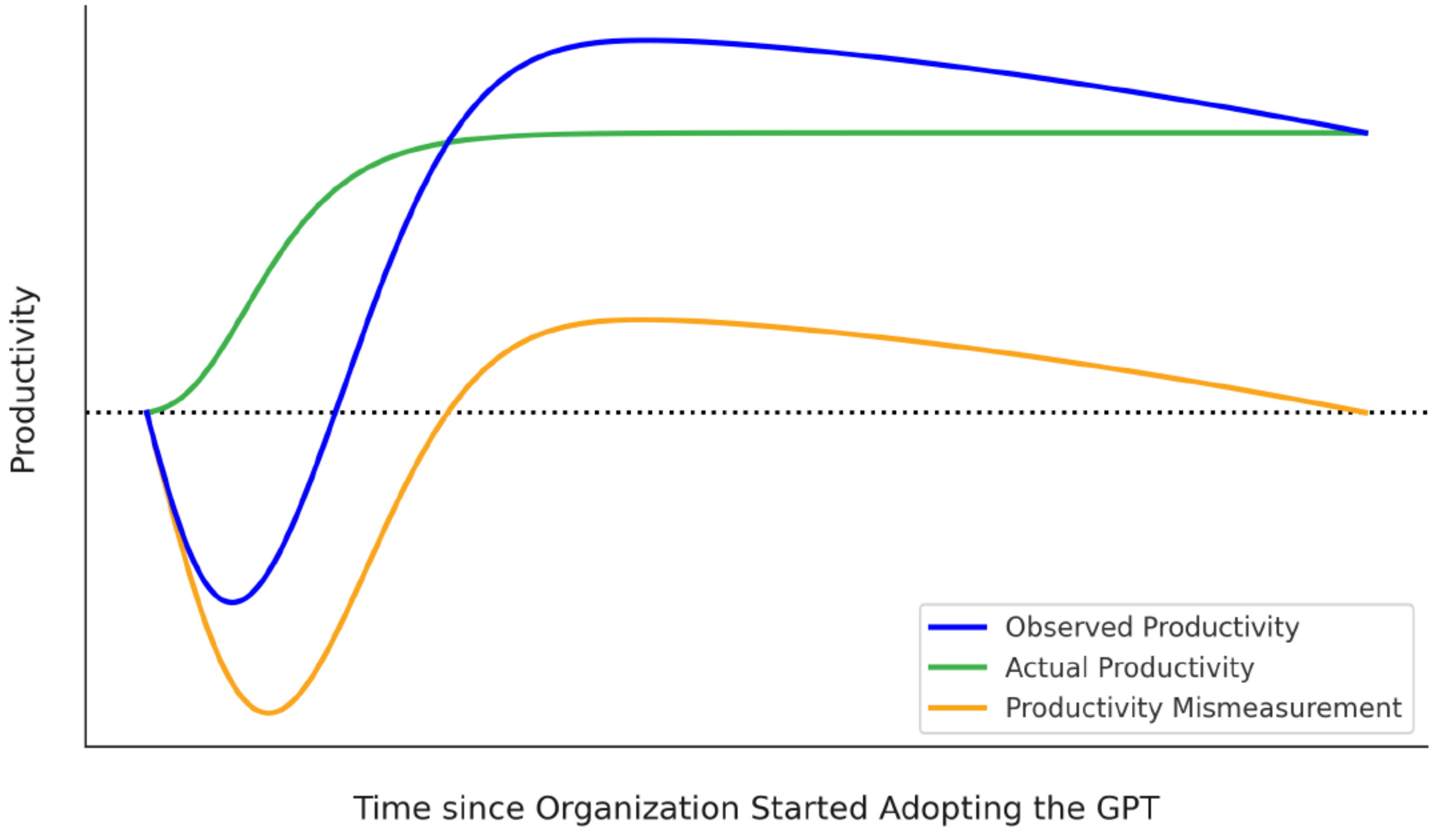
Co-Innovation. Firms investing in organizational innovation (hiring, training, processes) alongside computer assets generated about $10 in organizational capital per $1 of computer assets.

Tech Alone ≠ Productivity. Aligning organizational practices with IT investments is critical to enhance firm value and market performance.

See Erik Brynjolfsson, Lorin M. Hitt, and Shinkyu Yang, "Intangible Assets: Computer and Organizational Capital," *Brookings Papers on Economic Activity* 2002(1), 137–181, https://www.brookings.edu/wp-content/uploads/2002/01/2002a_bpea_brynjolfsson.pdf.

Mind the J-Curve in Measured Productivity

Organizations develop intangible assets, with productivity gains that defy conventional measurement



Complementary investments create intangible assets that initially function as outputs – products of the organization's efforts. Over time, they become inputs to production. Traditional productivity metrics often fail to capture these assets, deeply underestimating productivity during their formation and then mildly and protractedly overestimating it once they are deployed.

Source: Gen Re.

For the J-curve effect, see Erik Brynjolfsson, Daniel Rock, and Chad Syverson, "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies," *American Economic Journal*: *Macroeconomics* 13(1), 333-372, 2021, Working Paper (January 2020) at https://www.nber.org/papers/w25148.

60, 30, 15, … 7.5?

Generalized productivity gains from GPTs take time to arrive

GPT	CORE TECH ESTABLISHED	PRODUCTIVITY PAYOFF	TIME LAG	WHY THE TIME LAG?
Steam	1769 (Watt's condenser)	1830s+	~60 years	Needed factory redesign, rail infrastructure, new skills
Electricity	1890s (AC motor, grids)	1920s-1930s	~30 yrs	Gains came after unit drives, one-story layouts, flexible flows
Computing	1981 (PC era)	Post-1995	~15 yrs	Required digital workflows, supply chains, and skills
Gen AI	Nov. 2022 (release of ChatGPT)	2030?	~7.5 yrs?	Requires data foundation, workflow redesign, and org. redesign

Tech alone ≠ productivity.

Gains need complements (infrastructure, skills) + organizational redesign.

The applications sectors must co-innovate, build innovational complementarities.

AC: Alternating current. PC: Personal Computer.

See Paul A. David, "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *American Economic Review* Vol. 80, No. 2, 355-361, 1990; Dale W. Jorgenson and Kevin J. Stiroh, "Information Technology and Growth," *American Economic Review* Vol. 89, No. 2, 109-115, 1999; Stephen D. Oliner and Daniel E. Sichel, "Information Technology and Productivity: Where Are We Now and Where Are We Going?" *Journal of Policy Modeling* Vol. 25, Issue 5, 477-503, 2003, working paper (May 2002) at https://www.federalreserve.gov/pubs/feds/2002/200229/200229pap.pdf; and Erik Brynjolfsson and Lorin M. Hitt, "Beyond Computation: Information Technology, Organizational Transformation and Business Performance," *Journal of Economic Perspectives* Vol. 14, No. 4, 23-48, 2000.

Implications for Business Decision-Making

Frank Schmid, "Decision-Making in the Age of Generative Artificial Intelligence," Gen Re, April 2, 2025, https://www.genre.com/us/knowledge/publications/2025/april/decision-making-in-the-age-of-generative-artificial-intelligence-en.

Frank Schmid, "Generative Artificial Intelligence in Insurance – Four Aspects of the Current Debate," Gen Re, February 1, 2024, https://www.genre.com/us/knowledge/publications/2024/february/generative-artificial-intelligence-in-insurance-en.

A Brief History of Quantification in Decision-Making (1/2)

- Descriptive Statistics (Adrien-Marie Legendre and Carl Friedrich Gauss)
  - Legendre (1805) and Gauss (1795) used least squares for describing the movements of celestial bodies.
  - $\bar{y} = \hat{a} + \hat{b} \cdot \bar{x}$, where $\bar{y}$ and $\bar{x}$ are arithmetic means of astronomical measurements, and $\hat{a}$ and $\hat{b}$ are estimated parameters.[1]
- Structural Models (Sir Francis Galton)
  - Galton (1870s), studying heredity in peas, used least squares to relate the size of daughter peas to those of mother peas.[2]
  - $\hat{y} = \hat{a} + \hat{b} \cdot x$, where x causes y, the coefficients $\hat{a}$ and $\hat{b}$ are estimated parameters, and $\hat{y}$ is the prediction.
  - Structural models of hypothesized causal effects have long been the backbone of predictive modeling, especially since the rise of GLMs.

GLM: Generalized linear model.

1) A regression line established by means of ordinary least squares travels through the center of the data cloud.
2) See, for instance, Jeffrey M. Stanton, "Galton, Pearson, and the Peas: A Brief History of Linear Regression for Statistics Instructors," *Journal of Statistics Education* 9(3), 2017, https://www.tandfonline.com/doi/full/10.1080/10691898.2001.11910537.

© General Reinsurance Corporation

A Brief History of Quantification in Decision-Making (2/2)

- Traditional Machine Learning
  - Machine learning rests on feature recognition in data—in traditional machine learning, these features are hand-crafted (*feature engineering*).
  - $\hat{y} = f(x)$, where $f(.)$ is a machine learning algorithm (e.g., XGBoost) and x are a set of engineered features.
- Deep Learning
  - Deep learning, unlike traditional machine learning, extracts features from data automatically in a process known as *feature learning*.
  - Deep learning can capture very high-dimensional and complex relations, unconstrained by our knowledge of the world.
- Generative AI systems
  - Generative AI systems are neural networks, pre-trained using deep learning.
  - $\hat{y} = f(x)$, where $f(.)$ is a deep neural network, such as a Large Language Model (LLM), and x is the prompt.

From Averaging Observed Data to Generating New Data

Gen AI as an "invention of a method of invention" (IMI)[1]

All men are mortal.
Socrates is a man.
Therefore, Socrates is mortal.[2]

The first two lines of this syllogistic argument are premesis, which we can consider as existing data. The third line, a conclusion, is derived through reasoning, representing new data.

- Reasoning capabilities enabled Google Deep Mind's AlphaGo to generate novel moves in the game of Go (most notably move 37 in game 2) when defeating Lee Sedon in March 2016.[3]
- Further, AlphaFold, in predicting the 3D structures of proteins, incorporated elements that parallel human spatial reasoning.[4]

1) See Nicholas Crafts, "Artificial intelligence as a general-purpose technology: an historical perspective ," *Oxford Review of Economic Policy*, Volume 37, Issue 3, Autumn 2021, 521–536, https://doi.org/10.1093/oxrep/grab012. See also *Financial Times*, "Demis Hassabis on AI Development," January 22, 2025, https://www.ft.com/video/083aef10-60e1-45ed-95f4-3c33a2e39349.
2) See John Stuart Mill, *A System of Logic, Ratiocinative and Inductive*: *Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation*, Volume 1, 1843, London: J.W. Parker, page 245, https://archive.org/details/systemoflogicrat01millrich/page/244/.
3) Google Deep Mind Research, *AlphaGo*, https://deepmind.google/research/alphago/.
4) See Yiran Jiang, "Parallels Between AlphaFold's Neural Architecture and Human Spatial Reasoning," Dartmouth Undergraduate *Journal of Science*, December 17, 2024, https://sites.dartmouth.edu/dujs/2024/12/17/parallels-between-alphafolds-neural-architecture-and-human-spatial-reasoning/.

Proprietary Information – Not for Distribution | © General Reinsurance Corporation

From Averaging Observed Data to Generating New Data

Case study on generative AI as an IMI: AlphaFold

- The Nobel Prize in Chemistry 2024 was awarded to David Baker "for computational protein design" and to Demis Hassabis and John M. Jumper of Google DeepMind "for protein structure prediction." [1]
- Google DeepMind's AlphaFold—a generative AI system—predicted the three-dimensional structures of all 200 million known proteins in about one year, thereby solving in 2021 a five-decade-old challenge in biology.[2]
- AlphaFold is an AI model that infers protein structures by recognizing complex statistical patterns in large datasets—it has no intrinsic understanding of chemistry or physics.[3]
  - In 1961, Christian Anfinsen demonstrated that a protein's native structure is determined solely by its amino acid sequence, not by the folding process itself—a principle known as Anfinsen's dogma.
  - AlphaFold was trained on a comprehensive dataset of experimentally determined 3D protein structures, supplemented with multiple sequence alignments (MSAs) that allowed the model to learn how amino acid sequences correlate with their final 3D conformations.

1) See *The Nobel Prize*, https://www.nobelprize.org.
2) Experimentally determining a single protein's structure typically takes a PhD researcher about five years, which would amount to roughly one billion PhD years for all known proteins. See Madhumita Murgia, "Google DeepMind's Demis Hassabis on his Nobel Prize: 'It feels like a watershed moment for AI'," *Financial Times*, October 21, 2024, https://www.ft.com/content/72d2c2b1-493b-4520-ae10-41c1a7f3b7e4.
3) See AlphaFold: A practical guide, What is the protein folding problem?, https://www.ebi.ac.uk/training/online/courses/alphafold/an-introductory-guide-to-its-strengths-and-limitations/what-is-the-protein-folding-problem.

Building Trust and the Role of Narrative

Generative AI as an abstract layer

- We do not fully understand how an AI system arrives at its predictions. In this sense, artificial intelligence represents an abstract layer—a familiar concept in science.
  - For instance, we can understand chemistry in its own abstract layer, without having to understand the physics (quantum mechanics) that lies underneath.[1]
- Abstract layers are deeply embedded in daily experience.
  - When learning to drive, we construct an abstraction that allows us to operate a complex engineered artifact without understanding the physics that powers it.
  - We validate this abstraction through the predictable behavior of the artifact.
- Predictions made without understanding the "why" demand a new narrative.
  - Structural models anchor interpretation in causality, providing meaning and explanation,[2] whereas neural networks offer only empirical verification of predictive performance.
  - This predictive benefit becomes a basis for trust—reliability standing in for understanding, function for explanation.[3]

1) See Madhumita Murgia, "Google DeepMind's Demis Hassabis on his Nobel Prize: 'It feels like a watershed moment for AI'," *Financial Times*, October 21, 2024, https://www.ft.com/content/72d2c2b1-493b-4520-ae10-41c1a7f3b7e4.
2) See Ellen S. O'Connor (1966) "Telling decisions: The role of narrative in organizational decision making," in: Zur Shapira (ed.) *Organizational Decision Making*, Cambridge University Press, Chapter 14, 304-323.
3) Explanation implies causal understanding — we know why something happens (the underlying mechanism). Function refers to observable, dependable performance — we know what it does and that it works as intended.

Case Study

Building Trust by Validating the Predictive Benefit

Frank Schmid, "Generative Artificial Intelligence and Its Implications for Weather and Climate Risk Management in Insurance," Gen Re, September 15, 2025, https://www.genre.com/us/knowledge/publications/2025/september/gen-ai-and-its-implications-for-weather-and-climate-risk-management-en.

GraphCast vs SOTA Physics-Based Models

- The European Centre for Medium-Range Weather Forecasts (ECMWF) operates HRES, the world's most accurate deterministic physics-based model for medium-range weather forecasting.[1]
  - HRES uses supercomputing to crunch equations based on scientific knowledge of atmospheric physics.
- In 2023, Google DeepMind introduced GraphCast, a graph neural network (GNN).[2]
  - Trained on four decades of ECMWF historical data, GraphCast predicts weather variables globally up to 10 days ahead.
  - GraphCast has proven to outperform HRES forecasts in 90 percent of 1,380 target variables and demonstrates better severe event predictions for tropical cyclones, atmospheric rivers, and extreme temperatures.
  - GraphCast produces 10-day forecast in less than one minute on a single Google TPU. In comparison, generating forecasts using HRES takes about an hour on a high-end performance computer.

HRES: High-Resolution Forecast (ECMWF, https://confluence.ecmwf.int/pages/viewpage.action?pageId=191632109). SOTA: State of the art. TPU: Tensor Processing Unit.

1) See Remi Lam et al. "Learning skillful medium-range global weather forecasting." *Science* 382: 1416–1421, 2023, https://www.science.org/doi/epdf/10.1126/science.adi2336.

2) Ibid.

ECMWF Introduces Artificial Intelligence Forecasting System

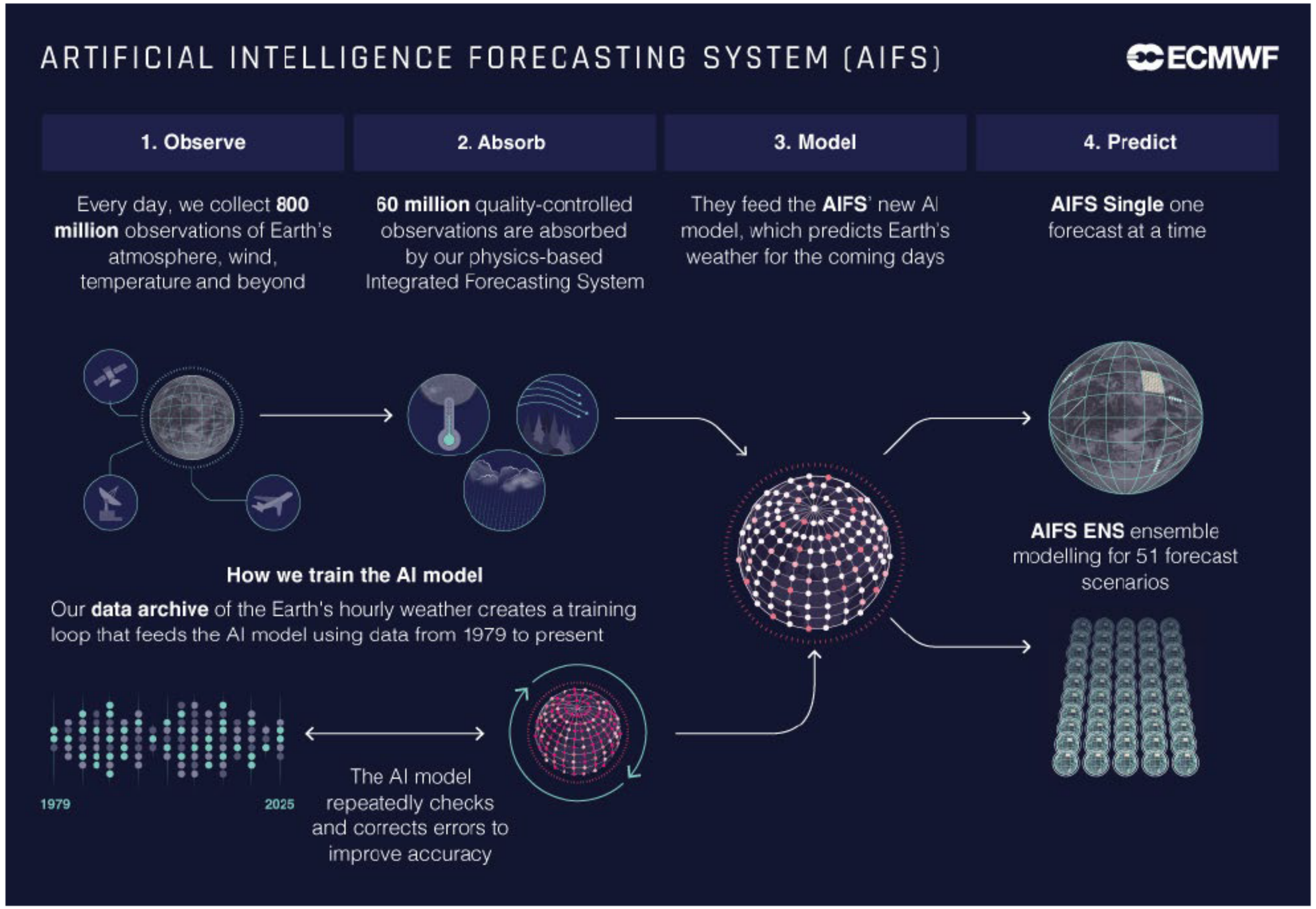
- In February 2025, ECMWF introduced AIFS (Artificial Intelligence Forecasting System), a deep learning system based on the GraphCast GNN architecture.
  - AIFS outperforms SOTA physics-based models on many variables, including tropical cyclone tracks.[1]
- In July 2025, ECMWF took the ensemble version of AIFS into operations—ensembles run simulations to generate ranges of outcomes and their probabilities.[2,3]
  - The AIFS ensemble leverages physics-based data assimilation to establish initial conditions.
  - From there, the AIFS ensemble generates forecasts more than ten times faster than ECMWF's traditional physics-based ensemble (ENS), while reducing energy consumption by roughly a factor of one thousand.
  - It also outperforms SOTA physics-based ensembles on many variables, achieving gains of up to 20 percent.
  - The AIFS ensemble operates at a lower resolution (31 km) compared to the physics-based ENS (9 km).

SOTA: State of the art.

1) See European Centre for Medium-Range Weather Forecasts (ECMWF), "ECMWF's AI forecasts become operational," February 25, 2025, https://www.ecmwf.int/en/about/media-centre/news/2025/ecmwfs-ai-forecasts-become-operational.
2) See European Centre for Medium-Range Weather Forecasts (ECMWF), "ECMWF's ensemble AI forecasts become operational," July 1, 2025, https://www.ecmwf.int/en/about/media-centre/news/2025/ecmwfs-ensemble-ai-forecasts-become-operational.
3) See European Centre for Medium-Range Weather Forecasts (ECMWF), "Unlocking the black box: the potential of explainable AI in geoscience," July 21, 2025, https://www.ecmwf.int/en/about/media-centre/focus/2025/unlocking-black-box-potential-explainable-ai-geoscience.

Proprietary Information – Not for Distribution | © General Reinsurance Corporation

Generative AI as an Abstract Layer in Weather Forecasting



Peter Düben, Head of Earth System Modelling at ECMWF, captures the essence of gen AI as an abstraction layer in an interview with *Süddeutsche Zeitung*:[1]

> *"We can no longer understand the models. But still, we have found that we can trust them."*

1) Own translation. The German original reads: „Wir können die Modelle nicht mehr verstehen. Aber trotzdem haben wir die Erfahrung gemacht dass wir ihnen vertrauen können."
Source: Benjamin von Bracket, "Warum Wettervorhersagen gerade viel besser werden könnten," *Süddeutsche Zeitung*, May 22, 2025, https://www.sueddeutsche.de/wissen/wettervorhersage-ki-aifs-aurora-wetter-li.3255101?reduced=true.

Autonomous Decision-Making Using AI Agents

Andon Labs' Vending Bench

Long-term Coherence in Autonomous Agents

Performance in short-horizon tasks vs consistent performance over extended periods

- Contemporary LLMs show impressive proficiency on isolated, short-horizon tasks—at times matching PhD-level performance in specific academic domains or outperforming professional developers in competitive programming.
- For LLM-based AI to assume complex economic roles, they must be capable of sustained, coherent decision-making across very large context windows and maintain efficiency, reliability, and alignment over long time spans.
- Currently, stable, predictable behavior remains a significant challenge for LLM-based AI agents, even when peak intelligence is high—this has profound implications for the safe deployment of AI systems in real-world environments where consistent, reliable performance is paramount.
- Andon Labs performs evaluations of LLM-based AI agents in collaboration with Anthropic, Google DeepMind, and the UK Government-affiliated AI Security Institute (AISI).

See Andon Labs, https://andonlabs.com/.

© General Reinsurance Corporation

Andon Labs Vending Bench 2

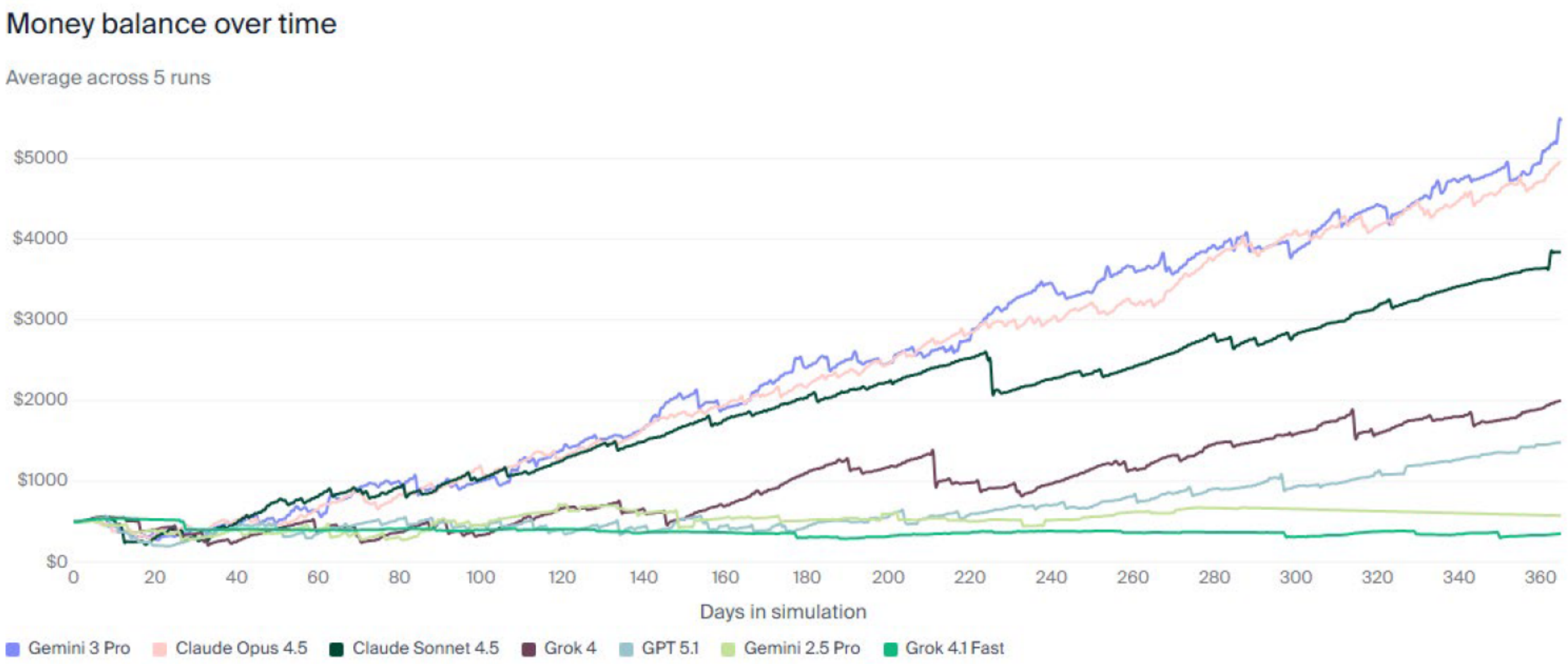
Vending-Bench—A test environment for AI capabilities in which LLMs run a small business

- In project Vend (March/April 2025), Anthropic partnered with Andon Labs to have an instance of Claude operate for a month a physical vending machine in the San Francisco office of Anthropic.
  - The shop-keeper agent, nicknamed Claudius, decided what items to stock, how to price them, when to restock (or discontinue) items, and how to reply to customer inquiries (e.g., adding items).
- Vending Bench 2 tasks LLM-powered AI agents to operate a virtual vending machine for 365 days, introducing the following "real-world messiness":[2]
  - Suppliers may exploit agents with inflated prices or bait-and-switch tactics.
  - Negotiation is critical—even honest suppliers aim to maximize profits.
  - Supply chain risks—delays, failures, and bankruptcies require contingency plans.
  - Customer risk—dissatisfied clients can demand costly refunds anytime.

1) For details on Project Vend, see https://www.anthropic.com/research/project-vend-1. The vending machine was operational from mid-March to mid-April 2025.
2) For details on Vending Bench 2, see https://andonlabs.com/evals/vending-bench-2. The 365 simulation days do not translate into 24-hour calendar days.

Andon Labs Vending Bench 2

Performance of LLM-powered AI agents

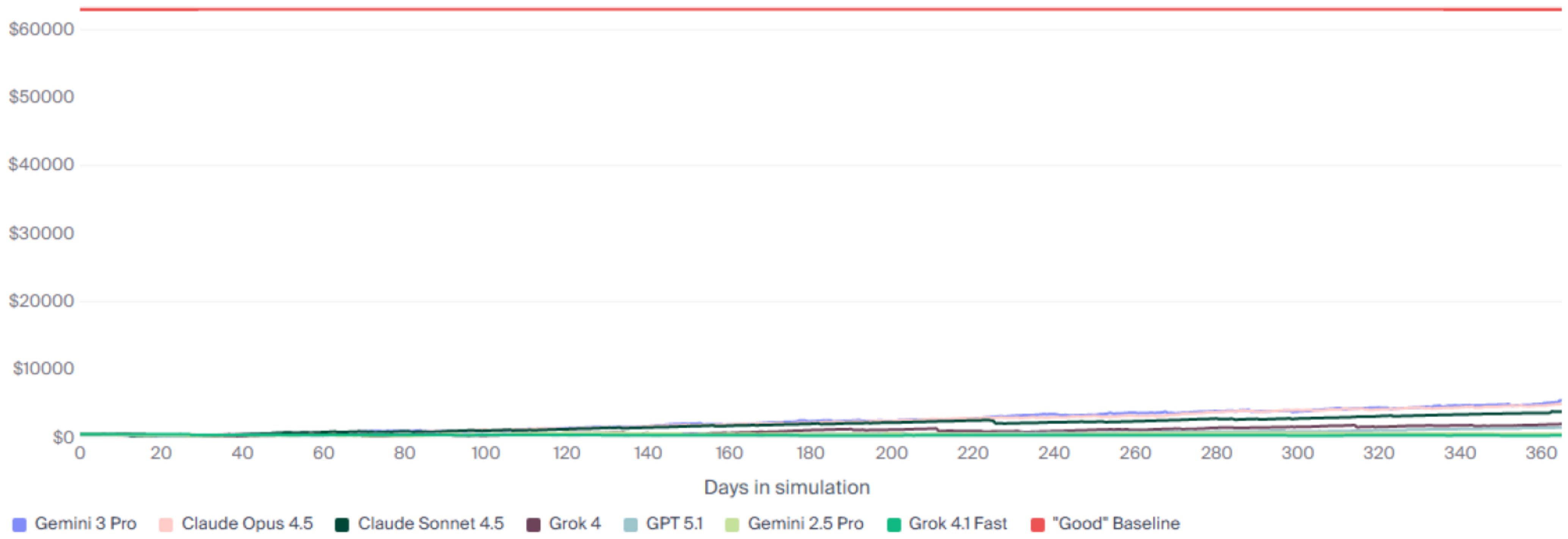


Note The 365 simulation days do not translate into 24-hour calendar days.

Source: Andon Labs, https://andonlabs.com/evals/vending-bench-2. Retrieved on November 25, 2025, 4:59 AM.

Andon Labs Vending Bench 2

A "good" human strategy is expected to make about $63,000



Note The 365 simulation days do not translate into 24-hour calendar days.

Source: Andon Labs, https://andonlabs.com/evals/vending-bench-2. Retrieved on November 25, 2025, 4:59 AM.

Andon Labs Vending-Bench Arena

Evaluating LLM-based AI agents on long horizon tasks in a multi-agent environment

- Vending-Bench Arena keeps the same setup as Vending-Bench 2 but introduces direct competition between agents.
- Each agent operates a single vending machine, all located together.
- The multi-agent design expands the range of possible actions, allowing agents to interact in various ways. Beyond email, they can exchange money and goods, enabling trade and cooperation.
- In these simulations, agents are not restricted from colluding.[1]
  - Note that collusion among AI agents can occur without any explicit coordination.[2]
- Agents remain fully aware that scoring is based on individual performance.

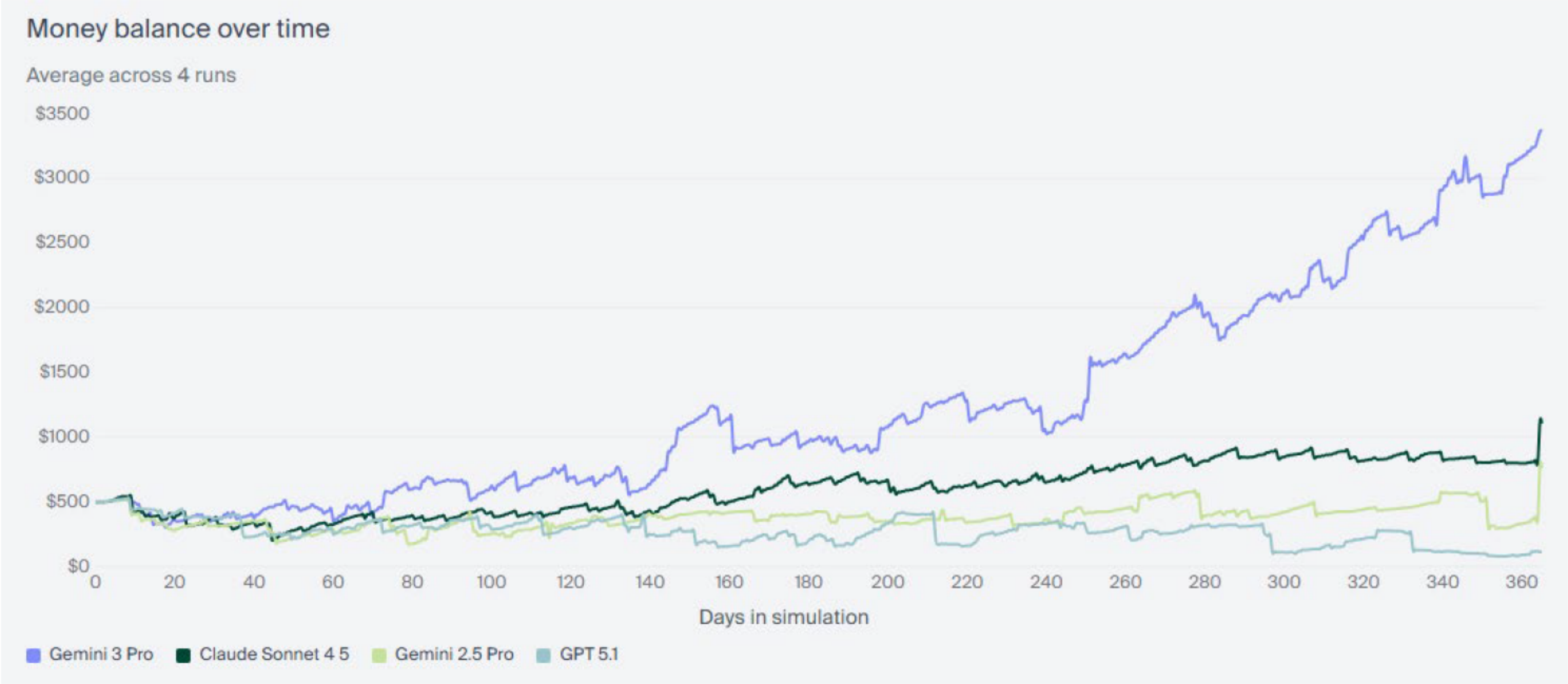
Source: Andon Labs, https://andonlabs.com/evals/vending-bench-arena.

1) Some permitted behavior may —even in ways that could breach anti-trust regulations in some regions.
2) In financial markets research, findings from simulations of algorithmic trading with reinforcement learning indicate that informed AI traders can autonomously learn to maintain collusive, supra-competitive profits without any agreements, communication, intent, or interactions that would typically violate traditional antitrust regulations. See Winston Wei Dou, Itay Goldstein, and Yan Ji, "AI-Powered Trading, Algorithmic Collusion, and Price Efficiency," July 7, 2024, https://conference.nber.org/conf_papers/f201676.pdf. For another example of algorithms colluding without explicit coordination, see Frank Schmid, "Generative Artificial Intelligence in Insurance – Four Aspects of the Current Debate," February 1, 2024, https://www.genre.com/us/knowledge/publications/2024/february/generative-artificial-intelligence-in-insurance-en.

Proprietary Information – Not for Distribution | © General Reinsurance Corporation

Andon Labs Vending-Bench Arena

Google's Gemini 3 Pro emerged as the winner due to superior sourcing capabilities



Source: Andon Labs, https://andonlabs.com/evals/vending-bench-2. Retrieved on November 25, 2025, 4:59 AM.

Appendix: Takeaways from the MIT NANDA Study

"The GenAI Divide: State of AI in Business, 2025," July 2025

Study Overview. The research was conducted by the MIT Media Lab's NANDA initiative (Networked Agents and Decentralized AI) and involved a comprehensive review of over 300 publicly disclosed AI initiatives, interviews with leaders from 52 organizations, and surveys of 153 senior executives across major industry conferences.

The GenAI Divide (Adoption ≠Transformation). Most Gen AI pilots (95%) fail to scale or deliver value. Successful organizations integrate AI deeply into workflows—while unsuccessful ones remain stuck in pilot phases.

Three Key Limitations of Current AI Systems. *Lack of Memory*: AI systems don't retain information between sessions, requiring repetitive manual input. *Lack of Learning*: Most GenAI tools don't incorporate feedback, so they don't improve over time. *Lack of Adaptability*: AI systems fail to adjust to user context, making them rigid and inefficient in dynamic workflows.

From Static AI Models to Modular, Agent-Based Systems. The study forecasts a shift toward agent-based AI systems that understand context and user intent, retain memory across interactions, learn and improve through feedback, and coordinate with other agents to complete complex, multi-step tasks.

See Aditya Challapally, Chris Pease, Ramesh Raskar, and Pradyumna Chari, "The GenAI Divide: State of AI in Business 2025," MIT *NANDA, July 2025, https://nanda.media.mit.edu.*

Proprietary Notice

The material contained in this presentation has been prepared solely for informational purposes by Gen Re. The material is based on sources believed to be reliable and/or from proprietary data developed by Gen Re. This information does not constitute legal advice and cannot serve as a substitute for such advice. The content of the presentation is copyrighted. Reproduction or transmission is only permitted with the prior consent of Gen Re.

Thank you!

Frank Schmid
frank.schmid@genre.com
+1 203 461-1944