RESEARCH REPORT

How much does AI increase productivity for R&D scientists in materials discovery?

541	47	3
CLAIMS	SOURCES READ	DISAGREEMENTS

PREPARED 2026-11-06
SESSION s_06GV5VYCPPFW5Z44

CONTENTS

Finding 3
Scope 4
Disagreement 5
Sources 6
Coverage 10
Receipt 10
Reproducing this 10

Finding

Written by a language model from the evidence in Sources. Every quote it draws on was checked against its source; the prose was not.

A preprint reported that AI-assisted teams discovered 44% more new materials and generated 17% more product innovations.[9,10,11,12,13] MIT says the paper's underlying data are unverified and that it cannot vouch for the study's truthfulness.[9,12,13] The claims therefore do not establish a verified percentage increase in materials R&D scientists' productivity.[9,12,13]

The preprint tracked a thousand materials scientists before and after an ML materials-generation tool was introduced at a large R&D company.[13] It reported little benefit for the bottom third of scientists, while the output of top researchers nearly doubled.[10] A separate theoretical model cites findings about materials-science gains from Toner-Rodgers (2024); it did not measure those gains itself.[22]

Materials-discovery output figures also conflict: one account says GNoME predicted 2.4 million stable materials, while another reports 2.2M crystal predictions, of which 381K were potentially stable.[18,27,28,11,15+9] Those figures describe predicted materials, while a later study found little evidence that GNoME had discovered novel materials.[25,26]

THIS REPORT IS INCOMPLETE

written from the 95 highest-ranked findings; 6 contradiction(s) withheld: not re-checked

Scope

What the question was decomposed into, and what became of each part.

1 How do government and peer-reviewed assessments characterize AI's current impact on materials R&D productivity, including limitations and caveats? RESEARCHED

2 What is the most-cited figure for AI-driven productivity or speedup in scientific/materials research, and how was it measured and has it been criticised? RESEARCHED

3 What evidence exists on AI productivity effects for R&D scientists generally (patents, publications, task completion), from studies or experiments? RESEARCHED

4 What quantitative productivity gains have been measured when AI/ML tools are used in materials discovery R&D (e.g. reduced experimental iterations, faster candidate screening, time saved)? RESEARCHED

5 SETTLING A DISAGREEMENT RESEARCHED
Google DeepMind's GNoME found 2.2 million crystals, of which 380,000 are predicted to be stable — or — Google DeepMind's GNoME predicted 2.4 million stable materials, expanding known stable materials by nearly 10×

6 SETTLING A DISAGREEMENT RESEARCHED
Google DeepMind's GNoME predicted 2.4 million stable materials, expanding known stable materials by nearly 10× — or — Google DeepMind's GNoME tool identified 2.2 million new stable crystalline materials, equal to roughly 800 years of

7 SETTLING A DISAGREEMENT RESEARCHED
GNoME predicted 2.2 million new materials — or — The document characterizes GNoME's output as several hundred thousand stable crystalline structures, not 2.2 million or

8 SETTLING A DISAGREEMENT NOT RUN
Google DeepMind's GNoME found 2.2 million crystals, of which 380,000 are predicted to be stable — or — The document states that Google DeepMind's GNoME tool identified 2.2 million new stable crystalline materials, equal to

9 SETTLING A DISAGREEMENT NOT RUN
The study was partly based on a survey of over 600 scientists in the US and UK carried out in July-August 2026 — or — The study collected 243 valid survey responses

10 SETTLING A DISAGREEMENT NOT RUN
A within-subject randomized controlled experiment involving GitHub Copilot found no statistically significant — or — GitHub claims a significant increase in speed of about 55% using Copilot in the auto-completion form

11 SETTLING A DISAGREEMENT NOT RUN
The 55% speed increase figure is a claim made by GitHub from its own customer data and evaluations, not an independent — or — A randomized controlled trial by Peng et al. (n=95) found a 56% speed increase for developers using GitHub Copilot

12	Which independent reviews or critiques dispute headline AI materials-discovery speedup figures, and what measurement problems do they identify?	RESEARCHED
13	What do randomised or controlled studies and field experiments measure on AI-assisted R&D productivity per scientist (effect sizes, hours saved, output per researcher)?	RESEARCHED
14	What do reviews and critiques identify as measurement problems behind headline AI materials-discovery or research-speedup figures?	RESEARCHED
15	What per-scientist effect sizes do randomised or controlled studies report for LLM/AI assistance on knowledge-work and R&D tasks, and how do they transfer to materials science?	RESEARCHED

4 of 15 sub-questions were not researched: this run reached its ceiling first. Nothing below rests on evidence they would have found.

Disagreement

Each block is one argument. Both sides are printed and neither is resolved — which of them to believe is the judgement this pack exists to inform, not to make. Read the quotes: two figures measured over different scopes are a difference, not a contradiction.

DISAGREEMENT 1 — 3 claims · 2 pairs

Google DeepMind's GNoME predicted 2.4 million stable materials, expanding known stable materials by nearly 10×.

cypris.ai/insights

> *Google DeepMind released GNoME (Graph Networks for Materials Exploration), predicting 2.4 million stable materials and expanding known stable materials by nearly 10×*

Google DeepMind's GNoME found 2.2 million crystals, of which 380,000 are predicted to be stable.

dataintelo.com/report

> *Google DeepMind's GNoME found 2.2 million crystals, of which 380,000 are predicted to be stable.*

1 more claim on this point is in the full evidence file.

DISAGREEMENT 2 — 2 claims

GNoME predicted 2.2 million new materials.

freethink.com/robots-ai

> *All told, GNoME predicted 2.2 million new materials.*

The document characterizes GNoME's output as several hundred thousand stable crystalline structures, not 2.2 million or 2.4 million new stable materials.

en.wikipedia.org/wiki

> *The tool proposes millions of materials previously unknown to chemistry, including several hundred thousand stable crystalline structures, of which 736 had been experimentally produced by the Massachusetts Institute of Technology, at the time of the release.*

DISAGREEMENT 3 2 claims

The document states the direct result as the discovery of 400,000 new stable crystals.

patmcguinness.substack.com/p

> *We woke up yesterday to find we just increased our knowledge of crystals many-fold, finding 400,000 new stable crystals, thanks to DeepMind's GNoME.*

DeepMind said 380,000 of the crystals were stable and were the best candidates for creating new materials.

siliconrepublic.com/machines

> *DeepMind said 380,000 of these crystals were "stable" and are the best candidates for creating new materials that could boost various forms of technology.*

Sources

Every source the prose cites, with the sentences mined from it. 541 claims were collected in total; all of them are in the evidence file you can download beside this document.

[1] https://arxiv.org/html/2507.18638v2
Rather than manipulating variables or conducting controlled experiments, the study focuses on naturally occurring behaviors and perceptions, which are best captured through survey methodologies.
The population in this study includes AI users from diverse backgrounds, primarily consisting of university students, educators, researchers, freelancers, content creators, and professionals in the fields of business, engineering, and information technology.
A total of 75.7% (184 out of 243) reported some level of agreement that AI tools help them complete tasks faster.

[2] https://www.researchprofessionalnews.com/rr-news-world-2026-9-ai-saves-scientists-time-but-adds-verification-tax-paper-finds
Almost three-quarters of respondents said AI saves them time, averaging 6.9 hours a week.
The paper says its findings suggest "AI holds significant potential to increase scientific productivity" while recognising they paint a "nuanced picture".
for 46 per cent of them, verification consumed more than a quarter of the time they had saved.

[3] https://campustechnology.com/articles/2026/09/23/ais-productivity-gains-in-science-tempered-by-time-spent-validating-outputs.aspx
Around three-quarters of surveyed researchers reported saving time through AI, and the average reported net saving was almost seven hours per week.
Nearly half of the surveyed scientists reported using some form of AI every day, and researchers who reported net time savings averaged just under seven hours a week.
The paper also reports that about 68% perceived greater access to insights from other disciplines, while roughly 65% said AI had increased the breadth of their research agendas.

[4] https://arxiv.org/html/2410.12944v2
One much-discussed study was a randomized controlled trial by Peng et al. [[12](#bib.bib12)] (n=95n=95), which found a 56% speed increase for developers using GitHub Copilot—an AI code assistant—, compared to those not using it.
We also found an interesting effect whereby developers who spend more hours on code-related activities per day were faster with AI.
Our best estimate of the size of this effect, controlling for factors known to influence developer time on task, stands at about 21%, although our confidence interval is large.

[5] https://mit-genai.pubpub.org/pub/v5iixksv
Peng et al. (2023) hired 95 professional programmers to develop an HTTP server in JavaScript. The group that used GitHub Copilot completed the task 55.8% faster, yet there was no effect on whether the task was completed.
For customer support via chat interfaces, Brynjolfsson, Li, and Raymond (2023) found that access to an AI-based conversational assistant increases the productivity of customer support agents (measured as issues resolved per hour) by 14% on average.
Programmers achieved a 27% increase in speed, whereas nonprogrammers reached performance levels comparable to experienced programmers with the assistance of GPT-3.

[6] https://www.mmi.ifi.lmu.de/pubdb/publications/pub/weber2024eics-llm/weber2024eics-llm.pdf

According to that, they claim a significant increase in speed of about 55% using Copilot in the auto-completion form [27].
The group with access to GitHub Copilot completed the task 55.8% faster.
In an additional self-assessment, the developers slightly underestimated the measured effect by suspecting a 35% increase in speed.

[7] https://www.science.org/doi/10.1126/science.adh2586
Participants with weaker skills benefited the most from ChatGPT, which carries policy implications for efforts to reduce productivity inequality through AI.
Inequality between workers decreased, and concern and excitement about AI temporarily rose.
We examined the productivity effects of a generative artificial intelligence (AI) technology, the assistive chatbot ChatGPT, in the context of midlevel professional writing tasks.

[8] https://www.oecd.org/en/blogs/2025/07/unlocking-productivity-with-generative-ai-evidence-from-experimental-studies.html
Recent experiments show that less-experienced or lower-skilled individuals tend to see the largest productivity gains when using generative AI tools.
In fact, more experienced ones can reap significant productivity gains as well. Their deeper knowledge, higher skill levels and greater contextual awareness can notably enable them to better interpret and apply AI-generated outputs.
At the same time, the impact of AI can vary significantly across industries, tasks, and individual roles.

[9] https://www.linkedin.com/posts/jonathan-usher_a-high-profile-ai-productivity-study-that-activity-7329679401372000256-0mrD
- 44% more new materials discovered.
- 17% more product innovations generated.
- 39% more patents filed.

[10] https://world.hey.com/ian.mulvany/data-showing-ai-productivity-gains-in-materials-science-0a2825b4
AI-assisted researchers discover 44% more materials, resulting in a 39% increase in patent filings and a 17% rise in downstream product innovation.
This preprint - Artificial Intelligence, Scientific Discovery, and Product Innovation - <https://aidantr.github.io/files/AI_innovation.pdf> from Aidan Toner Rogers at MIT <https://economics.mit.edu/people/phd-students/aidan-toner-rodgers> indicates that AI may be having a beneficial impact in materials science research.*
while the bottom third of scientists see little benefit, the output of top researchers nearly doubles.

[11] https://market.us/report/ai-in-materials-discovery-market
44% higher material discovery rates are achieved when researchers use AI tools.
Patent output rises by 39% for AI-assisted research teams.
Independent researchers have already synthesized 736 of these predicted materials.

[12] https://www.lesswrong.com/posts/M2GzdAGbxwinERSEt/a-widely-shared-ai-productivity-paper-was-retracted-is
The headline results was that the AI caused a 44% increase in materials discovery at the firm, with a productivity increase of 81% for top-decile scientists.
Both Acemoglu and MIT have [publicly withdrawn support from the paper](https://www.wsj.com/tech/ai/mit-says-it-no-longer-stands-behind-students-ai-research-paper-11434092?st=qsvacG&reflink=desktopwebshare_permalink) and urged that it be retracted.
MIT has no confidence in the provenance, reliability or validity of the data and has no confidence in the veracity of the research contained in the paper.

[13] https://forum.effectivealtruism.org/posts/YwaJxLEZkFtdzDCeD/a-widely-shared-ai-productivity-paper-was-retracted-is
The headline results was that the AI caused a 44% increase in materials discovery at the firm, with a productivity increase of 81% for top-decile scientists.
MIT has no confidence in the provenance, reliability or validity of the data and has no confidence in the veracity of the research contained in the paper.
It claimed the productivity of a thousand material scientists was tracked before and after the introduction of an machine learning material generation tool.

[14] https://en.wikipedia.org/wiki/Google_DeepMind
In November 2023, Google DeepMind announced an Open Source Graph Network for Materials Exploration (GNoME).
However, according to [Anthony Cheetham](https://en.wikipedia.org/wiki/Anthony_Cheetham "Anthony Cheetham"), GNoME did not make "a useful, practical contribution to the experimental materials scientists."

[15] https://www.instagram.com/p/DfGB9kuD6l1
Built by Google DeepMind and published in Nature, GNoME works like a seasoned chemist with infinite patience.
Source: Google DeepMind / Nature, GNoME materials discovery 2023
Trained on known crystal structures, it proposes candidate materials, tests them computationally, learns from failures, and iterates — millions of times.

[16] https://deepmind.google/blog/millions-of-new-materials-discovered-with-deep-learning
External researchers in labs around the world have independently created 736 of these new structures experimentally in concurrent work.
Computational approaches drawing from the Materials Project, Open Quantum Materials Database and WBM database boosted this number to 48,000 stable crystals.
Using materials from the Materials Project and insights on stability from GNoME, the autonomous lab created new recipes for crystal structures and successfully synthesized more than 41 new materials

[17] https://www.freethink.com/robots-ai/google-ai-discovers-2-2-million-new-materials
To date, researchers have independently produced only 736 of them in a lab.
GNoME trained on data available from the Materials Project and, using those 48,000 previously discovered inorganic crystals as a foundation, began constructing theoretical crystal structures.
Over the last decade, researchers have used computational methods to up that number to 48,000.

[18] https://time.com/6340681/deepmind-gnome-ai-materials
To test whether the materials GNoME predicts to be stable are indeed stable, Google DeepMind partnered with external researchers who successfully synthesized 736 of them.
GNoME was trained using data on material structures and their stability from the Materials Project.
Although the materials will still need to be synthesized and tested, steps which can take months or even years, the latest development is expected to accelerate the discovery of new materials

[19] https://hiwavemakers.com/blog/ai-materials-discovery-funnel-gnome-736-real
External researchers in labs around the world independently created 736 of them.
Berkeley Lab's autonomous A-Lab, working with DeepMind researchers, synthesized more than 41 new materials using AI-guided robotic techniques.
The bottleneck is not computation. It is synthesis: figuring out a sequence of real-world steps that produces the predicted arrangement of atoms.

[20] https://pubmed.ncbi.nlm.nih.gov/40711807
This review highlights the recent transformative impact of artificial intelligence (AI), machine learning (ML), and deep learning (DL) on materials science, emphasizing their applications in materials discovery, development, and optimization.
AI-driven methods have revolutionized materials discovery through structure generation, property prediction, high-throughput (HT) screening, and computational design while advancing development with improved characterization and autonomous experimentation.
Publication type: Journal Article, Review

[21] https://www.gov.uk/government/publications/ai-for-science-strategy/ai-for-science-strategy
Investment and attention are being drawn to the idea that increasing the productivity of scientific research will be the most valuable application of AI.
The hardest steps require breakthroughs moving beyond simulated or narrowly controlled domains, real-world validation of the efficacy of AI-designed therapeutics and materials, and reproducible evidence that AI can accelerate discovery in a way that tangibly benefits scientists and citizens.
Of particular importance is understanding how scientists will adapt to using new tools that can hallucinate or make errors, and ensuring the adoption of AI does not negatively impact research integrity.

[22] https://arxiv.org/html/2604.02189v1
[Toner-Rodgers (2024)](#bib.bib19) finds that the benefits of AI in materials science are strongly heterogeneous, with substantial gains concentrated among scientists with good judgment capabilities, while others struggle to assess the quality of AI-generated suggestions.
AI increases the probability that a given recombination attempt succeeds, effectively shrinking the distance between ideas and enabling more ambitious research strategies
the effect of increasing the share of AI-automated R&D tasks is non-monotonic: firms initially target more radical innovations, but beyond a threshold of human-AI complementarity, they shift the focus toward incremental innovations.

[23] https://www.alphaxiv.org/abs/2410.12944
We also found an interesting effect whereby developers who spend more hours on code-related activities per day were faster with AI.
Our best estimate of the size of this effect, controlling for factors known to influence developer time on task, stands at about 21%, although our confidence interval is large.
While they found a large interaction effect suggesting that high-intensity coders might benefit more from AI ($a\ \beta=-0.29$ interaction term), this result did not reach the threshold for statistical significance ($p=0.018$ for the whole model, but the specific interaction term was not significant at $p<0.05$).

[24] https://community.materialssquare.com/t/google-puts-the-world-in-check-with-their-ai-2-2-million-new-materials-discovered/672
Despite the vast potential, many of these materials remain to be synthesized.
an AI tool that discovered over 2 million potential new inorganic crystal compositions

The AI dramatically accelerates material discovery by predicting stable structures using deep learning.

[25] https://web.archive.org/web/20240822195305/https://www.siliconrepublic.com/machines/deepmind-ai-study-criticism-materials-discovery
A new study analysing a recent claim from Google-owned DeepMind suggests there is little evidence the company managed to find novel materials using AI.
But a group of researchers have taken issue with some of the claims made by DeepMind and claim they found "scant evidence" for compounds that "fulfil the trifecta of novelty, credibility, and utility".
"Polymers, glasses, metal-organic frameworks, heterostructures and composites are only a few excluded materials classes that come to mind and that are each infinite in their scope," the researchers said.

[26] https://www.siliconrepublic.com/machines/deepmind-ai-study-criticism-materials-discovery
But a group of researchers have taken issue with some of the claims made by DeepMind and claim they found "scant evidence" for compounds that "fulfil the trifecta of novelty, credibility, and utility".
DeepMind said 380,000 of these crystals were "stable" and are the best candidates for creating new materials that could boost various forms of technology.

[27] https://cypris.ai/insights/ai-accelerated-materials-discovery-in-2025-how-generative-models-graph-neural-networks-and-autonomous-labs-are-transforming-r-d
Google DeepMind released GNoME (Graph Networks for Materials Exploration), predicting 2.4 million stable materials and expanding known stable materials by nearly 10×
While specific commercial implementations are often proprietary, several indicators point to widespread industrial adoption.

[28] https://patmcguinness.substack.com/p/deepminds-gnome-discovers-new-materials
They have just increased the total number of known stable crystals by an order of magnitude in one announcement.
improving the efficiency of materials discovery by an order of magnitude
Using it, they were able to predict 2.2 million new crystals.

[29] https://pdpspectra.com/blog/ai-materials-discovery-gnome-2026
reported synthesizing 41 new compounds out of 58 targets over 17 days
And inorganic crystals exclude polymers, glasses, metal-organic frameworks, composites — most of what industry actually deploys.
*Its output is a list of structures that *should* be stable according to DFT.*

[30] https://medium.com/@sabine_vdl/results-from-recent-experimentation-of-the-effects-of-ai-on-knowledge-worker-productivity-insights-e12fa058a0b1
For instance, performance reviews of 758 consultants by [Harvard Business School](https://www.hbs.edu/faculty/Pages/item.aspx?num=64700) revealed that AI tools could significantly elevate the quality of work.
Within the consulting industry, these advancements have shown that workers using AI support complete tasks more expeditiously.

[31] https://www.linkedin.com/posts/digitalcio_ai-coding-is-massively-overhyped-report-activity-7378588384144953344-1fxx
But those claims appear to be massively overblown, as The Register reports, with researchers finding that productivity gains are modest at best — and at worst, that AI can actually slow down human developers.
"Surprisingly, we find that when developers use AI tools, they take 19 percent longer than without — AI makes them slower," the researchers wrote.

[32] https://dataintelo.com/report/materials-discovery-ai-market
Small experimental datasets remain a technical limit, which is why NIMS and four Japanese chemical makers developed methods to predict properties from only a small number of experiments.
Google DeepMind's GNoME found 2.2 million crystals, of which 380,000 are predicted to be stable.

[33] https://metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study
When developers are allowed to use AI tools, they take 19% longer to complete issues—a significant slowdown that goes against developer beliefs and expert forecasts.

[34] https://www.linkedin.com/posts/kaseylobaugh_an-autonomous-laboratory-for-the-accelerated-activity-7373518279446851584-5Bfz
In practice, that means scientists generated as much useful data in weeks as traditional labs would in months or years.
DeepMind's GNoME: 2.2M crystal predictions, 381K potentially stable materials.

[35] https://cmr.berkeley.edu/2025/10/seven-myths-about-ai-and-productivity-what-the-evidence-really-says
While not specific to Generative AI, the findings of this review indicate that when decision-support systems are highly but not perfectly reliable, users become over-trusting, leading to a 12 percent increase in commission errors (e.g., accepting incorrect AI suggestions) and slower detection of rare anomalies.
While not specific to Generative AI, the findings of this review indicate that when decision-support systems are highly but not perfectly reliable, users become over-trusting, leading to a 12 percent increase in commission errors

Yet, a systematic review of 74 studies on automation bias and complacency documents a concerning pattern (Goddard et al., 2012). While not specific to Generative AI, the findings of this review indicate that when decision-support systems are highly but not perfectly reliable, users become over-trusting, leading to a 12 percent increase in commission errors (e.g., accepting incorrect AI suggestions) and slower detection of rare anomalies.

[36] https://www.schrodinger.com/life-science/resources/webinar/accelerating-materials-discovery-with-physics-informed-ai-ml
Artificial Intelligence (AI) and machine learning (ML) are reshaping materials science, accelerating the discovery of novel materials and optimizing formulations with unprecedented speed and precision.

[37] https://www.alphaxiv.org/abs/2508.06642
Most significantly, AF showed a clear tendency to increase with the dimensionality of the parameter space.
The simulations using five 2D Gaussian functions with identical Lipschitz complexity but varying contrast demonstrated that maximum EF increases monotonically with contrast.

[38] https://news.stanford.edu/stories/2026/04/digital-chores-productivity-boost-research
is one of the first to examine this largely overlooked aspect of the AI revolution

[39] https://www.fhi.mpg.de/2177332/2025-12-11_AI-Driven-Materials-Discovery
These natural deviations are often overlooked by current AI systems, leading to inaccurate predictions about how a material will behave once synthesized.
The study demonstrates that current AI and simulation workflows often assume perfect crystal lattices.
While AI and simulation-based discovery promised to accelerate this process, the new findings show that ignoring structural disorder results in unreliable predictions.

[40] https://www.bloorresearch.com/is-ai-productivity-a-myth-fusion-economy
Goldman Sachs, on 2025 US productivity data

[41] https://neurips.cc/virtual/2025/129013
Existing benchmarks for computational materials discovery primarily evaluate static predictive tasks or isolated computational sub-tasks.

Coverage

SOURCES READ	47
CONSIDERED	96
NOT REACHED	49
COULD NOT READ	0

Considered counts every distinct source this session had in front of it: read, refused, and passed over once the source limit was reached.

Receipt

QUOTES RE-READ	541
CONFIRMED	540
NOT FOUND IN THE SOURCE	0
MOVED WITHIN THE SOURCE	1
SOURCE UNREACHABLE	0
SKIPPED	0

Every quote is machine-verified against its source. Whether the sources are right is your judgement.

Reproducing this

This pack was assembled from one research session. The session is kept, so the same question can be re-run later and the answers compared.

SESSION	s_06GV5VYCPPFW5Z44
RESEARCHED	2026-11-06

PACK BUILT 2026-11-08