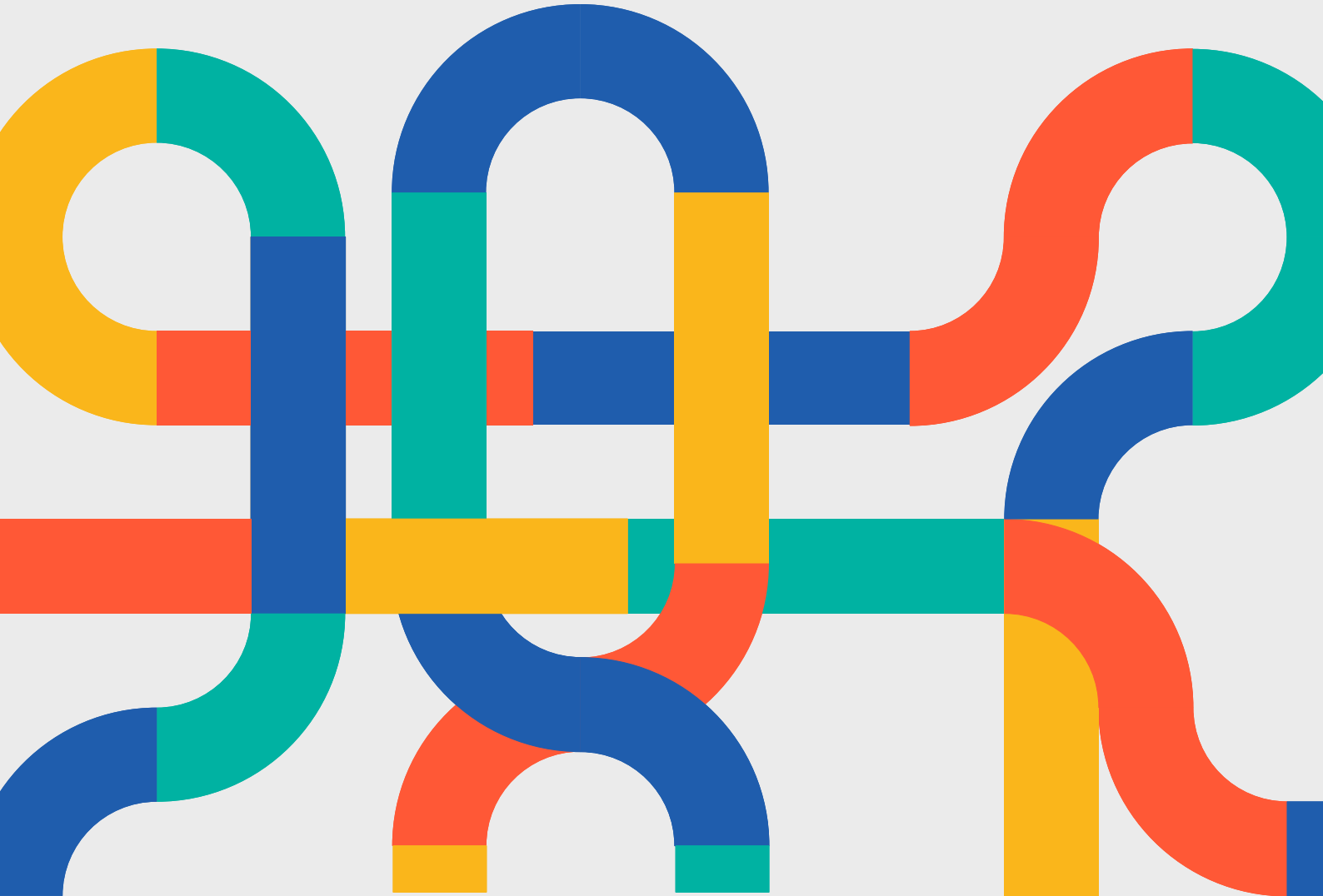
Innovation Growth Lab

AI in Science: Emerging evidence from AlphaFold 2

Summary Report

David Ampudia Vicente and George Richardson

25th November 2025



Contents

Executive Summary

Artificial intelligence tools and methods are diffusing within scientific research. Questions exist over their effects on the productivity, diversification and quality of research, however, robust evidence of systemic impacts of AI in science are limited.

AlphaFold 2 is an AI tool, developed by Google DeepMind in 2021, that addresses a longstanding problem in structural biology - protein structure prediction. The performance of AlphaFold 2 was unanticipated and is recognised for addressing this challenge. The tool and a database of 200 million predicted protein structures have been released for free. These characteristics make AlphaFold 2 a valuable case study.

Our analysis studies 5 million academic publications, clinical articles, patents and protein structures to investigate the impact of AlphaFold 2. We measure the association of AlphaFold 2 with scientific productivity, novel research and applied outcomes. We assess direct and indirect impacts of the tool against a baseline of structural biology research, and other high impact, contemporary developments in four key areas.

Scientific reach: 550,000 publications are linked to AlphaFold 2 directly and indirectly, involving almost 2 million unique researchers, with the latter measure of reach exceeding that of other frontier developments. However, AlphaFold 2's influence continues to grow, while others have plateaued. We also estimate that 218,000 articles incorporate elements of AlphaFold 2 into their methodology.

Experimental structural biology: Researchers building on AlphaFold 2 demonstrate a 45% - 49% increased rate of experimental protein structure submissions, higher than rates for those adopting other frontier developments. Influence of the tool is also associated with consistently more unique protein structures, demonstrating a tendency to increase novelty. However, structure resolution is lower in these less chartered parts of the protein universe.

Academic productivity and quality: Links to AlphaFold 2 lead to a modest increase in publication rates for researchers (2.5%) and laboratories (5.1%), similar to other frontier developments. Citation counts for research papers building on AlphaFold 2 and other frontier developments exhibit increases between 25% and 30%. Normalising citation counts by field and year still yields positive associations, with AlphaFold 2 performing strongly for researchers and laboratories.

Applied research and innovation: There is a generally positive impact associated with AlphaFold 2 across the areas that we examine, consistent with an emerging technology, and similar to more established frontier developments. In particular, individual researchers using AlphaFold 2 have a 9.3% greater likelihood of publishing disease related research. Papers linked to AlphaFold 2 are twice as likely to be cited in clinical articles, double the increase of other AI protein prediction developments, but researchers themselves see no increase. We find that research and researchers building on AlphaFold 2 are between 22.6% and 34.2% more likely to be cited by a patent, similar to other frontier developments.

Conclusions

AlphaFold 2 is generally associated with positive impacts across the dimensions investigated here. The effect sizes are typically on par with, or exceed, other frontier developments in terms of magnitude, and are more consistent, particularly with respect to other frontier AI developments. In addition, it is now leveraged over a much larger population of researchers. This demonstrates that high accuracy, predictive AI tools can diffuse across science as quickly as high impact research built on more established approaches.

We also find that AlphaFold 2's impacts on protein structure exploration hold significance for AI in science. First, the tool's high association with novel protein structures highlights that AI tools which are used to make predictions to map large knowledge spaces can encourage greater exploration of previously uncharted areas. This may be as a result of lowering the associated risks and costs for researchers. Second, its strong positive association with experimental outputs is indicative of good integration within teams and existing methods and domain knowledge, challenging the notion that AI and existing research methods do not mix well. This is reinforced by AlphaFold 2's emerging association with applied outputs, which suggests that despite its recency, it is capable of translational impact that spans fundamental and applied research.

Our study suggests that when AI is applied to a bottleneck problem with a clear definition and existing data, it can unlock progress in science. However, our study is not exhaustive. We are not able to verify that AI does not lead to 'streetlight effects' (systemic concentration on AI shaped problems), or that it is reaching its full potential. We suggest nonetheless that the positive effects should be capitalised upon, while science policy and funding decisions should be used to ensure science uses AI in ways that are complementary to existing methods of knowledge generation, and that additional tools and infrastructure for researchers can be developed to bolster impacts.

The work presented in this report is supported by a comprehensive literature review, methodology and results paper available from the Innovation Growth Lab.

1. Introduction

Artificial intelligence (AI) has become increasingly integrated into scientific research, driven by advancements in neural networks and the collection of large scale data. Its rapid adoption is evident across all fields; research publications referencing AI keywords rose from 5% in 2020 to over 8% in 2022, with approximately half of this activity now occurring outside of traditional computer science and mathematics disciplines (Gargiulo *et al.*, 2022; Duede *et al.*, 2024). This integration has fueled discussions about AI's potential to accelerate discovery and widen the scope of scientific exploration (Wang *et al.*, 2023; Agrawal, McHale and Oettl, 2024; Bail, 2024; Messeri and Crockett, 2024; Sumner, 2024). However, the exact nature of its impact remains an open empirical question, with debates centering on whether AI will ultimately unlock substantial productivity gains and enrich science or, conversely, create "streetlight effects" that narrow research focus toward data-rich, low-hanging fruit (Tranchero *et al.*, 2022).

The present study contributes to this evidence base through an in-depth, quantitative analysis of AlphaFold 2, a protein structure prediction system released in 2021 (Jumper *et al.*, 2021). AlphaFold 2 represents a unique, high-impact case study as it addressed a critical, long-standing bottleneck in structural biology and the wider life sciences. Its successful development earned its creators the 2024 Nobel Prize in Chemistry (Callaway, 2020; Brzezinski *et al.*, 2024). By comparing AlphaFold 2's impact against other contemporary frontier developments in structural biology (both AI-intensive and traditional) we investigate its influence on experimental discovery, research productivity, and translational outcomes. The following sections summarize the existing literature on AI's impact in science before introducing additional details about the development and release of AlphaFold 2.

1.1 The Impact of AI on Science

Productivity

One primary promise of AI in science is its capacity for automated, high-throughput prediction, leading to increased efficiency across the research process. Early application of AI in science involved researchers building bespoke models using open-source packages (Pedregosa *et al.*, 2011; Abadi *et al.*, 2016; Paszke *et al.*, 2019). Today, there are an increasing number of large pre-trained models (e.g., MetNet for forecasting, GNoME for material properties) and general-purpose generative AI (ChatGPT, Gemini, Claude) that assist with tasks ranging from code writing to simulating social science experiments (Sønderby *et al.*, 2020; Andrychowicz *et al.*, 2023; Jakubik *et al.*, 2023; Merchant *et al.*, 2023; Vaswani *et al.*, 2023).

Emerging evidence suggests a positive association between AI and high-impact research outputs. Studies indicate that AI-related papers are more likely to be classified as 'hit'

papers, ranking in the top 5% of citations within their field and receiving an average of 10.32% more citations than peers (Bianchini, Müller and Pelletier, 2022; Gao and Wang, 2024). While this may indicate that AI is enabling higher-quality work, it could also be partially driven by the visibility of AI as a trending research topic, underscoring the need for careful comparative analysis.

Scientific Creativity

Beyond efficiency gains, AI holds the potential to enhance scientific creativity through its ability to map unchartered parts of large, combinatorial knowledge spaces. In addition, it can enable "inverse design", where scientists specify a desired functional property (for example, for a new chemical synthesis) and an AI tool generates and prioritizes candidates for real-world experimentation (Schweidtmann *et al.*, 2018; Savage *et al.*, 2024). AI could act as a general method of invention, accelerating discovery by efficiently directing resources to unknown regions of the knowledge map (Agrawal, McHale and Oettl, 2018). Agentic workflows, or "co-pilots," that decompose problems and simulate or automate R&D steps are a potential next step (Hammond, 2023).

However, this potential is weighed against concerns that over-reliance on AI could diminish scientific diversity. Critics suggest AI may lead to monocultures by diverting research focus toward problems best suited for computational methods and potentially hinder the development of foundational theories (Krenn *et al.*, 2022; Messeri and Crockett, 2024). While empirical studies in narrow domains show AI-enabled methods are more likely to traverse less-explored problem spaces (Chenthamarakshan *et al.*, 2023), wider evidence suggests a lack of integration between AI and non-AI work, with deep learning papers in health sciences showing lower recombinatorial novelty (Bianchini, Müller and Pelletier, 2022; Duede *et al.*, 2024).

Complementarities with other capabilities and domain knowledge

The integration of AI into R&D could change the demand for scientific skills (Bianchini, Müller and Pelletier, 2022). While some tasks may be automated or deskilled, leading to a decreased demand for certain skills, new roles combining domain knowledge with AI expertise are likely to emerge, similar to other domains (Brynjolfsson and Mitchell, 2017; Brynjolfsson, Li and Raymond, 2023; Eloundou *et al.*, 2023). AI tools might also have the capacity to lower the barrier to entry for researchers, enabling newcomers or labs with fewer resources to participate in frontier research (Wang *et al.*, 2023).

Conversely, there are concerns that AI could exacerbate existing research inequalities. The development and effective application of AI systems still relies on access to high-performance computing resources and established knowledge networks. For instance, pharmaceutical firms with greater in-house domain knowledge are more effective at capitalizing on AI-driven discoveries (Tranchero, 2024). Successfully adopting AI is often predicated on collaboration between AI experts and domain-specific researchers, meaning labs with these connections have an advantage (Bianchini, Müller

and Pelletier, 2024). Therefore, while AI offers the theoretical potential for democratization, success currently depends on existing capacities, resources, and institutional connections.

1.2 AlphaFold 2: A Case Study for AI in Science

Proteins are a fundamental building block for life, and their complex three-dimensional structures, dictated by the sequence of amino acids, determine their function. Experimentally determining these structures, traditionally through methods like X-ray crystallography (a breakthrough that earned the 1962 Nobel Prize), has historically been a challenging, time-consuming process often requiring several years of labour (Hill and Stein, 2020). For several decades, the structural biology community also sought computational methods to predict protein structures and accelerate this process. The Critical Assessment of Structure Prediction (CASP) competition has been held since 1994 to benchmark results.

The field reached an inflection point with AlphaFold, an AI tool developed by Google DeepMind. After achieving strong results in CASP13 (2018), its successor, AlphaFold 2, achieved an accuracy comparable to that of experimental methods at CASP14 (2020) (Callaway, 2020). This achievement effectively "solved" the decades-old protein folding problem (AlQuraishi, 2020; Callaway, 2020; Perrakis and Sixma, 2021; Bertoline *et al.*, 2023; Elofsson, 2023; Brzezinski *et al.*, 2024). Google DeepMind subsequently released the AlphaFold 2 source code and, in partnership with EMBL-EBI, released a database of over 200 million freely downloadable predicted protein structures (Jumper *et al.*, 2021; Varadi *et al.*, 2022).

The unexpected success of AlphaFold 2 and its release represents a somewhat exogenous shock to the field, one that could have significant downstream impacts, including understanding disease mechanisms, drug discovery, and vaccine development (Duran-Frigola, Mosca and Aloy, 2013; Hazra and Patra, 2021; Higgins, 2021; Saplakoglu, 2024). Early quantitative analysis suggests that, while it has not increased the volume of papers published, its use is associated with an 8% increase in citations for adopting authors and enables the study of longer, more complex, and novel proteins (Yu, 2024).

Our study seeks to expand on this initial evidence by systematically comparing its impact on scientific outputs and translational research against similar high-impact innovations.

2. Methodology

2.1 Principles

We set out to answer two main research questions:

1. Does AlphaFold 2 lead to impactful research outputs?
2. Do these effects differ from those seen with other frontier developments?

Our study examines the impact of AlphaFold 2 in science and innovation at a systemic level, adhering to a number of criteria that inform our methodological design:

1. **Study several units of analysis:** Study the impacts on researchers who build on AlphaFold 2 (and other frontier developments), and on their publications.
2. **Broad coverage of R&I outputs:** Examine a range of research outputs from protein structures, to academic publications, to clinical trials and patents.
3. **Track direct and indirect impacts:** Consider research that builds on AlphaFold 2 through direct citations, and diffuse influence through indirect citation chains.
4. **Comparative analysis:** We compare AlphaFold 2 to leading contemporary developments in structural biology with similar usage characteristics and to a 'business-as-usual' baseline of a wide body of structural biology research.
5. **Longitudinal analysis:** We examine trends and changes after the adoption of AlphaFold 2, accounting for pre and post-treatment characteristics to isolate effects.

2.2 Data

To achieve the criteria for our analysis, we draw on a range of data sources:

- **OpenAlex:** An open database of academic papers, authors and their metadata.
- **Semantic Scholar:** A database of publications that provides enriched citation data.
- **Protein Data Bank:** The primary global database of verified protein structures.
- **UniProtKB:** A large database of protein sequence and function information.
- **Document Object Identifiers:** Standard IDs to link publications and other entities.
- **iCite:** A publication database which categorises clinical articles.
- **Medical Subject Headings:** A taxonomy of concepts related to medical science.
- **The Lens:** An open database that links patents to academic publications.
- **Cooperative Patent Classification:** A taxonomy of patent topics and technologies.

2.3 Methods

Our analytical approach to map AlphaFold 2's impact on structural biology, and its diffusion into applied research, relies on three main components: a data collection and enrichment; researcher and laboratory identification; defining counterfactual frontier developments; and a stringent regression specification.

Data collection and enrichment

The process begins by identifying three core AlphaFold 2 publications: Jumper et al. (2021), Evans et al. (2022), and Varadi et al. (2022). We then construct a broad citation network around these works, collecting a large corpus of papers that extends beyond structural biology. This network allows us to distinguish between research directly linked to the core papers, termed "adjacent"' and work linked indirectly through secondary citations, termed "downstream". To capture the nature of this influence, we further classify these citation chains based on citation intent, distinguishing between references that provide background context and those that indicate a direct methodological influence.

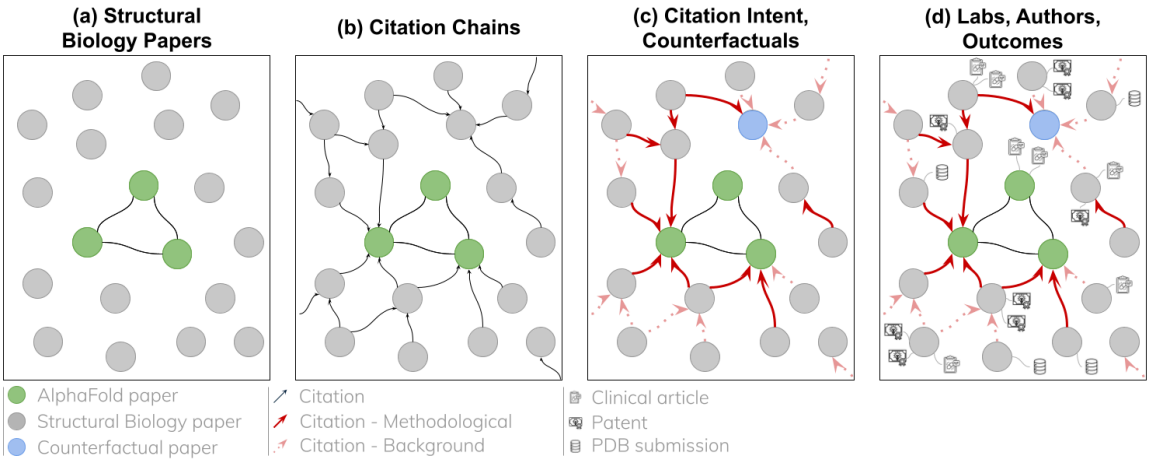


Figure 1. Overview of the dataset construction pipeline. It illustrates the multi-stage process for building our dataset, designed to track AlphaFold 2's impact from its core publications to adjacent and downstream research, as well as to applications outside the immediate research sphere.

We then establish a comparative baseline by identifying other frontier developments in structural biology, a selection process outlined in Figure 2. These developments are identified based on high citation counts and a distribution of methodological versus background citations that is comparable to AlphaFold 2. We classify these into three distinct counterfactual categories: AI-intensive protein prediction, non-AI protein prediction, and other structural biology innovations.

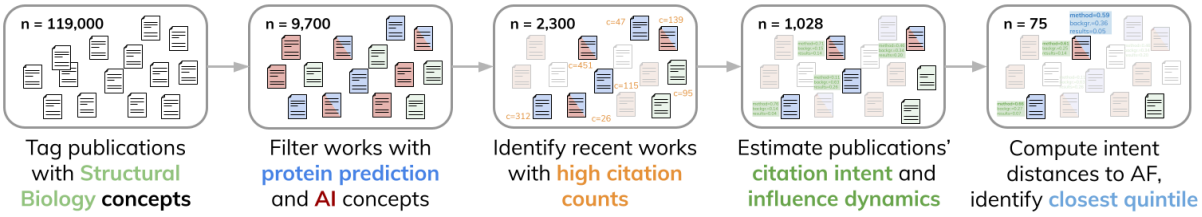


Figure 2. Selection process for identifying relevant papers in structural biology using OpenAlex concepts, CWTS topics, and citation intent metrics from Semantic Scholar.

Once the core and counterfactual networks are established, we enrich the dataset with extensive metadata. This includes linking publications to tangible outcomes such as patents, clinical trials, and protein structure submissions. Using author and publication metadata, we identify established researchers and laboratory leads, specifically targeting those with consistent and continuous senior authorship. Finally, to assess the quality and novelty of the scientific output, we incorporate several metrics describing the protein structures associated with publications in our dataset, such as structure similarity scores.

Regression analysis

We analyse the impact that building on AlphaFold 2 or other frontier developments has on research and innovation outcomes for publications, researchers and laboratories. By

comparing to the structural biology baseline, we obtain results that permit analysis of AlphaFold and the frontier research categories independently and relative to each other.

For example, in the first set of results we show in this report, we show the likelihood that a publication will be linked to an experimentally verified protein structure if it is linked to AlphaFold 2 or one of the frontier development counterfactuals, relative to the likelihood for papers that build on the structural biology baseline. In parallel, we also show the effects on the likelihood of having work linked to protein structure submissions for researchers and laboratories.

Our regression analysis seeks to identify associations between AlphaFold 2 use and impacts on outcome measures, while accounting for potential confounding factors and selection effects, controlling for unobserved fixed effects that might influence research outcomes, and controlling for observables like the primary field of a publication.

For researchers, we include additional constraints and methods, including Coarsened Exact Matching, to mitigate against differences in pre-treatment characteristics among affecting outcomes. We also carry out initial analysis to verify that pre-treatment career trajectories among our comparison groups are similar.

A full description of our data sources, dataset construction and regression methodology is available in the comprehensive literature review, methodology, and results paper, linked in the Executive Summary.

3. Findings

3.1 Scientific Reach

Breakthrough scientific tools are typically characterised by overcoming a longstanding bottleneck and then achieving widespread diffusion among researchers. AlphaFold 2 satisfies the first dimension, and we can assess the second by measuring uptake and influence, quantifying how, where and how quickly it has been integrated into research.

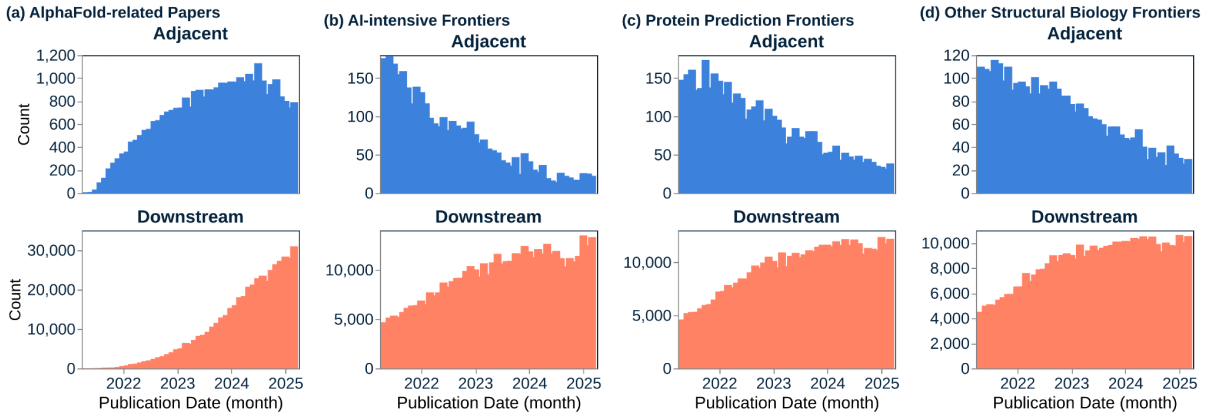


Figure 3. Monthly counts of new adjacent and downstream publications across frontier groups.

We find nearly 41,000 papers that are adjacent to the AlphaFold 2 core papers, and 640,000 downstream papers. Around 68% of adjacent papers cite AlphaFold 2 methodologically, and 38% in downstream research. This suggests that there is a substantial body of work that is strongly influenced by AlphaFold 2. The papers linked to AlphaFold 2 make up 16% of all publications in our final sample of structural biology work. We also observe that the number of new publications influenced by AlphaFold 2 continues to grow on a monthly basis, while the number of new papers linked to the other frontier developments has plateaued since 2024, and now show linear cumulative growth. This is seen clearly in Figure 3.

We estimate that AlphaFold 2 has been built on methodologically by 63,000 unique researchers in adjacent papers, and over 726,000 in downstream works. The latter figure is approaching the number of baseline structural biology papers in our sample (those not linked to counterfactual frontier developments).

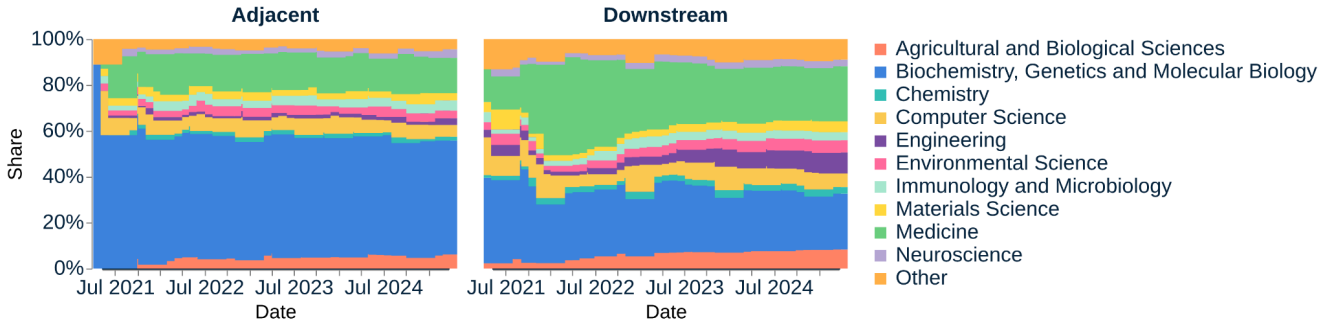


Figure 4. Monthly distribution of AlphaFold 2-related publications by primary topic.

Using the primary field associated with each publication, we investigate disciplinary trends, with trends over time being shown in Figure 4. We find that more than half of papers adjacent to AlphaFold 2 are focused on biochemistry. In downstream publications, this number is 30%, with medical sciences comprising a share that is similar in size, demonstrating diffusion into applied research.

3.2 Protein Structures

AlphaFold 2 has the potential to expedite protein structure determination, and to support exploration of proteins with structures that were previously very unfamiliar. In this sense, it is important to measure its impact on the number of proteins being characterised, and on the direction of research taken by scientists.

Our analysis focuses on proteins whose structures have been determined experimentally, rather than structures with only a predicted structure. We do this because the data provide a reliable measure of activity, and because real-world verification of a protein structure is the ultimate goal. This creates a level playing field with other frontier developments.

In our descriptive analysis of protein structures, shown in Figure 5, we notice that on average, more established methods tend to be implicated in protein structure discoveries

that are less novel, while AlphaFold 2 is consistently associated with more novel structures. Other frontier methods are associated with higher structural novelty using some metrics, but negative results also occur. We also note some differences between laboratory and researcher outcomes, with researchers experiencing stronger effects (although some results show low significance). Full results are shown in Figure 6.

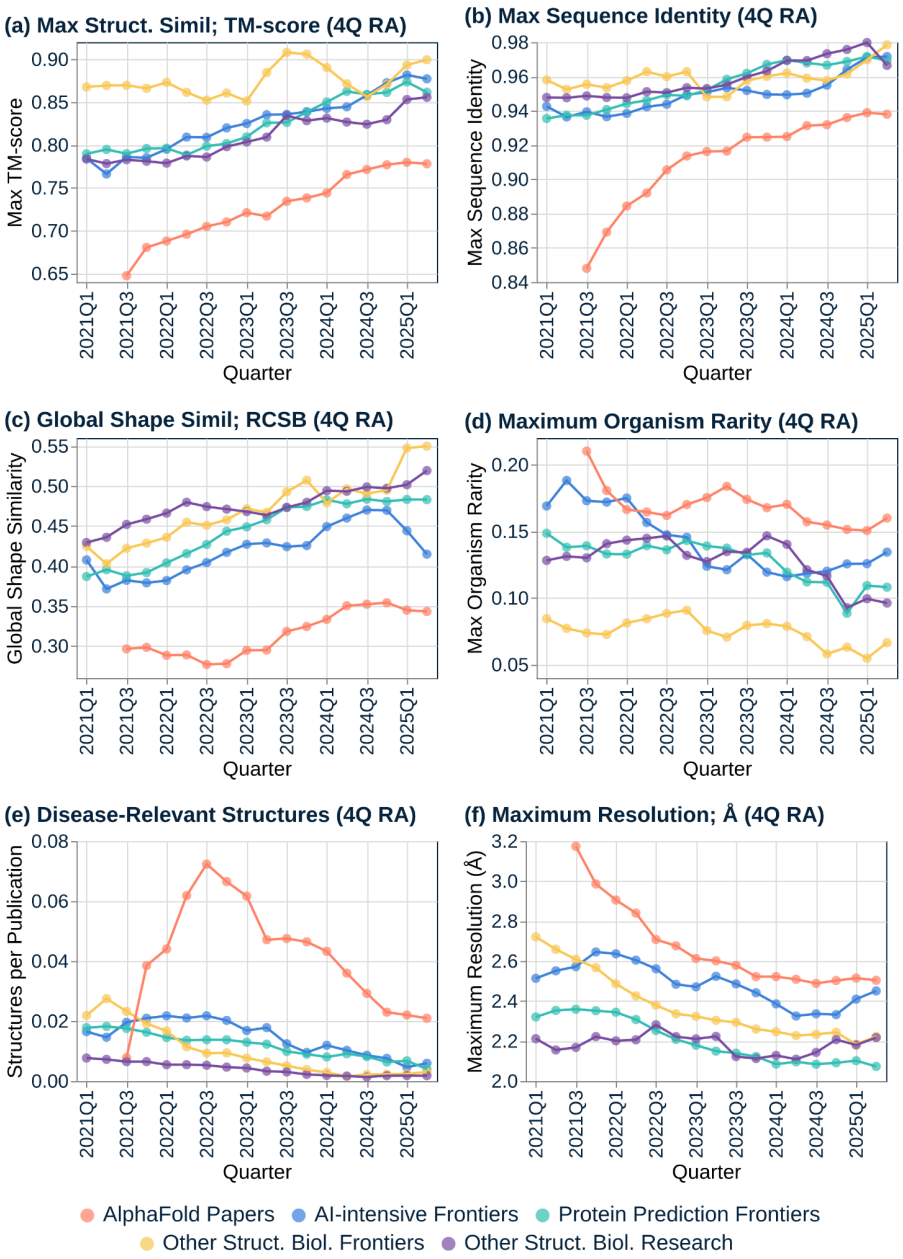


Figure 5. Quarterly rolling mean similarity scores, organism rarity, disease relevance, and resolutions for papers linked to AlphaFold 2 and other frontier developments.

We separately draw attention to one distinct measure of novelty: the likelihood of research involving disease-relevant structures. This metric indicates the direction of research, rather than abstract protein structure novelty. It is also a measure that exhibits strong and significant effects with a distinct dynamic. First, AlphaFold 2 appears to be associated

with a 39% increase in disease related research activity for laboratories, paralleled by other frontier structural biology developments. However, AlphaFold 2 is uniquely associated with a 137% increase in disease related research for established researchers.

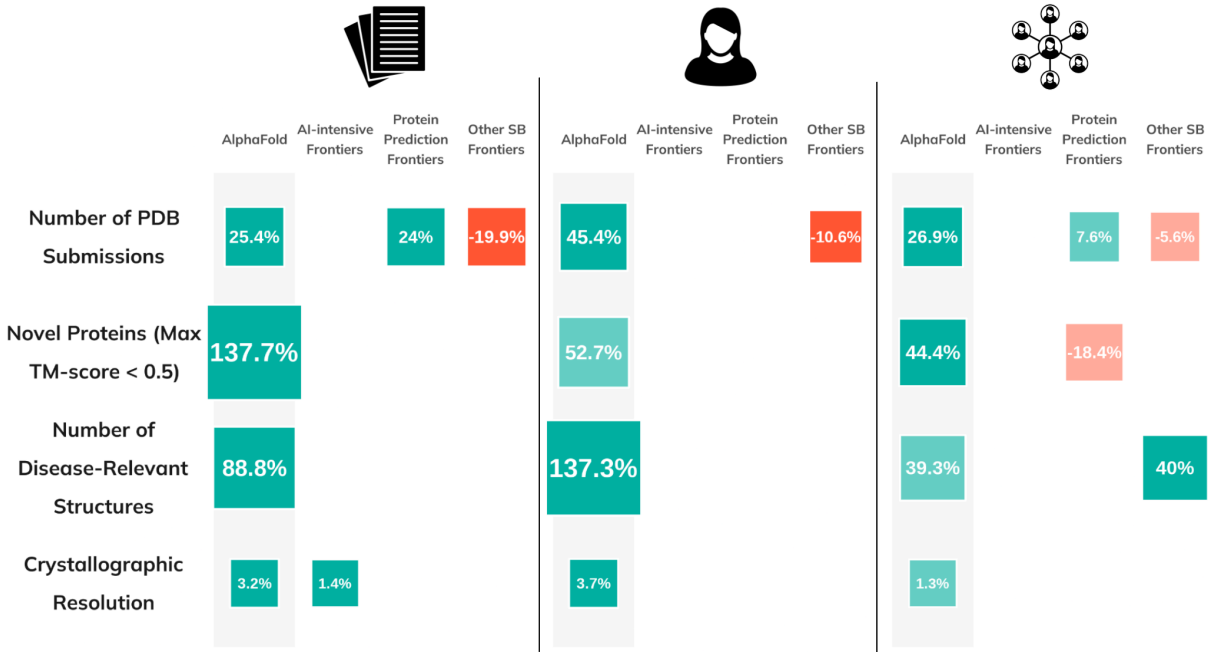


Figure 6. Association between AlphaFold 2, other frontier developments and PDB submissions plus key protein structure metrics. Coloured squares show effect magnitudes; darker shades indicate higher statistical significance. Empty cells (no square) mark effects that are statistically insignificant.

This pattern of results can be interpreted in several ways:

- The accuracy of protein structure predictions from AlphaFold 2 compared to other prediction methods reduces the risk of exploring less characterised areas of the protein space and protein structures that may be harder to determine.
- A significant number of researchers choose to make the trade off between investigating more novel proteins and a loss in experimental resolution that also characterises these structures.
- Comparable impacts experienced by individual researchers and laboratories building on AlphaFold 2 may be in part due to predictions reducing the complexity of the team-based science historically required for protein structure determination.
- Experienced researchers, who may become laboratory leads in future, are focusing more on protein structures implicated in disease related research, suggesting potential future trends towards more combined novel protein structure determination *and* more applied research agendas.

3.3 Academic Output

One speculation surrounding AI is that it will significantly increase the productivity of research, however empirical evidence for this impact is limited. In addition there are conflicting opinions about the potential impact of AI on the quality of science. AI might

support scientists to focus on more promising areas of scientific discovery, raising overall research quality, or lead to a slew of formulaic and incremental publications from researchers who are incentivised to treat AI outputs themselves as new knowledge.

We measure academic output in two ways. First, we attempt to gauge the effects of AlphaFold 2 and other frontier developments on researcher publication volumes, to investigate aggregate effects on scientific productivity. Second, we investigate the count of citations received by those researchers and their papers, as a proxy for research impact.

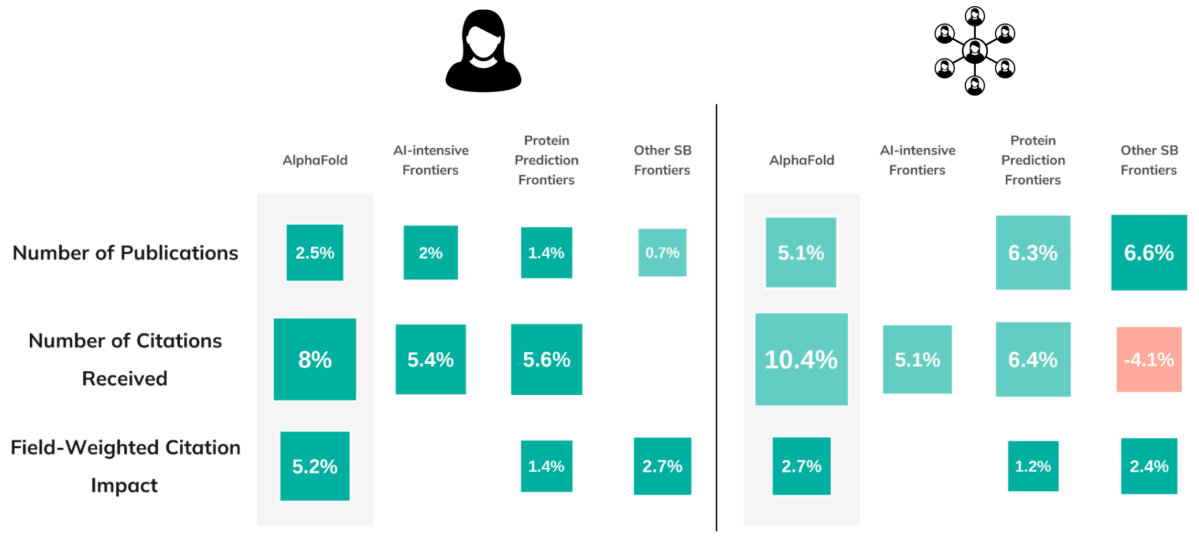


Figure 7. Association between AlphaFold 2, other frontier developments and PDB submissions plus key protein structure metrics. Coloured squares show effect magnitudes; darker shades indicate higher statistical significance. Empty cells (no square) mark effects that are statistically insignificant.

Figure 7 shows the full results of our regression analysis. For publication volumes, the results are modest. Researchers building on AlphaFold 2 gain a 2.5% increase in publication output, while a marginally significant effect is detected for laboratories at 5.1%. All other frontier developments show positive impacts of a similar magnitude (bar non-AI frontier structural biology developments for researchers and other AI frontier developments for laboratories). The high level pattern holds when only methodological citations are accounted for, with laboratories gaining the highest benefits from AlphaFold 2, and researchers gaining slightly more from other AI methods and protein prediction methods.

By design, our citation impact analysis suggests that all paper outputs drawing on frontier developments gain an average and similar increase in citation counts. For AlphaFold 2, this increase is 28.9%. This pattern holds for both citation counts and the field weighted citation index (FWCI), which takes into account field- and year-specific citation patterns.

The impact on citation counts for individual researchers and laboratories are smaller in size, but exhibit an interesting pattern. When only raw citation counts are taken into account, researchers and laboratories gain a comparable citation count benefit from AlphaFold 2 (8% and 10.4% respectively), surpassing the positive but smaller associations observed for most other frontier developments. When FWCI is used to account for field

differences, a distinct pattern emerges: AlphaFold 2 yields robust positive associations for both researchers and laboratories, whereas other AI frontier developments show no significant association with citation impact.

We repeat the analysis to compare methodological and background citations and note similar patterns, although there is limited difference between the two types of use. In most cases, the associations linked with methodological use of AlphaFold 2 and frontier developments are similar to background use.

There are several ways to interpret the results here, which suggest a range of possible dynamics at play:

- Despite our attempts to account for them, it is not possible to entirely disentangle pre-treatment trends and selection effects, meaning that we cannot say with certainty whether researchers, who appear more productive and prolific in this study, are adopting frontier developments, including AlphaFold 2, or whether those tools and techniques are themselves driving the pace and quality of research.
- Nonetheless, the citation results are somewhat unsurprising. We would expect any scientific works building on frontier developments in their field to result in greater downstream citation impact, due to their recognition for field advancement.
- We know that the publications downstream of the source papers in this study are spread across a range of fields. We also know that many of the citations for AlphaFold and AI frontier are in computer science. The gains in FWCI associated with AlphaFold 2 compared to other AI frontier developments could be explained by AlphaFold 2 having wider applicability in work outside of its original intended application.

3.4 Applied Research

While the expectations around the impacts of AI on scientific knowledge generation are contested, it is clear that AlphaFold 2 is able to map large numbers of candidates for downstream applications in the bioscience and biomedical domains. This use of AI to break bottlenecks in search and optimisation problems is a pattern common to research efforts on AI in many disciplines. It is therefore important to assess the translational impacts of AlphaFold 2.

We measure the influence of AlphaFold 2 and frontier developments in structural biology across three dimensions of applied activity: disease-focused research activity, clinical trials, and patenting. This provides insight into the influence of AlphaFold 2 and comparable frontier developments in structural biology beyond academic discovery.

For scale, we present the number of translational outputs associated with AlphaFold 2 papers in Table 1.

	Count	
Article type	**Adjacent**	**Downstream**
Disease related publication	3,097	76,320
Clinical publication	239	12,787
Patent	100	3,800

Table 1. Counts of adjacent and downstream applied articles linked to AlphaFold 2.

Disease relevance is measured by the presence of Medical Subject Headings (MeSH) C-class terms in publications. This branch of the MeSH taxonomy is titled 'Diseases'. By this definition, we find that 9.4% of papers adjacent to AlphaFold 2 and 14.8% of downstream papers are disease related, as seen in Figure 8. This is lower than general structural biology work, of which around 24% is disease related, but on average higher than other frontier AI developments (10.6% and 12.2% for adjacent and downstream papers respectively).

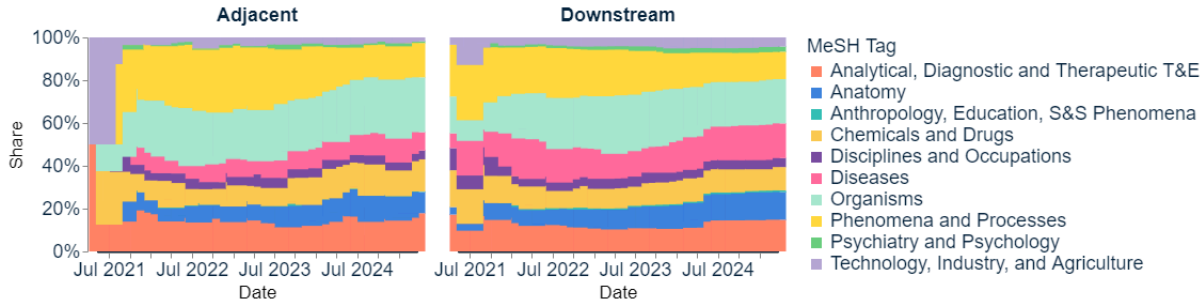


Figure 8. Monthly distribution of AF2-related publications by MeSH category.

In our regression analysis, we find that papers building on AlphaFold 2 and other frontier AI are no more likely to be disease related than structural biology in general, while works that build structural biology and protein structure prediction frontier are positively associated with MeSH C terms. These associations are shown in Figure 9. This relationship becomes positive for AlphaFold 2 when we observe the work of researchers (9.3%) and laboratories (5.0%).

Clinical articles are identified through their study design and content, covering randomised controlled trials, clinical trials, and observational studies that directly involve human subjects or clinical specimens. We find that papers citing AlphaFold 2 have a positive association with citations from clinical work, with a likelihood almost double that of the structural biology baseline papers. Positive effects are also seen for AI and protein prediction frontier papers, albeit at just over half this magnitude. However, at the researcher and laboratory level, we see no statistically significant association between AlphaFold 2, or any other frontier developments, and clinical citations (bar a modest positive association for protein prediction frontier developments built on by laboratories).

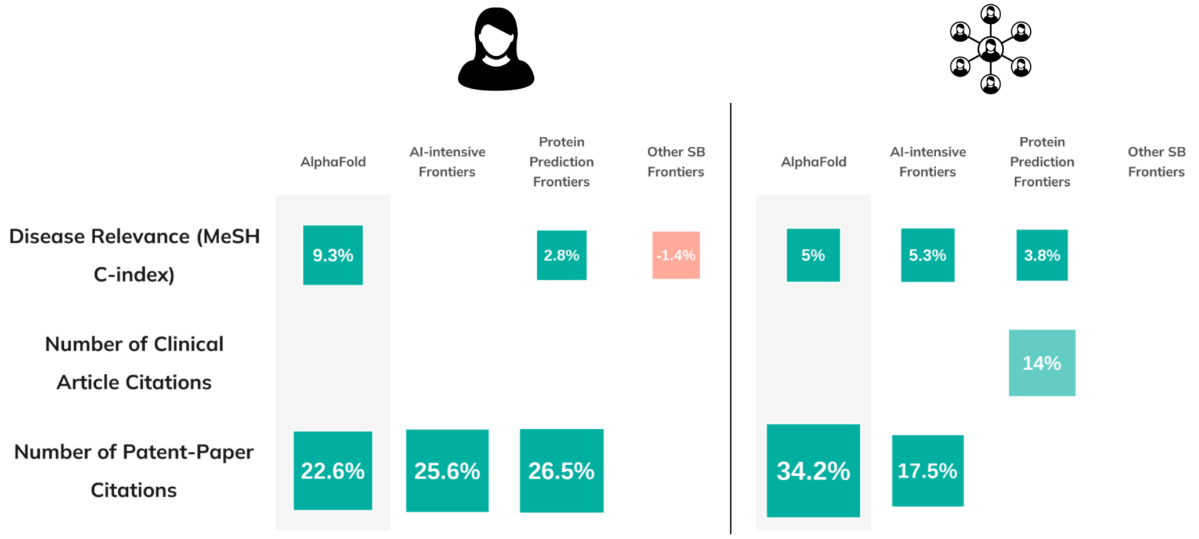


Figure 9. Association between AlphaFold 2, other frontier developments and PDB submissions plus key protein structure metrics. Coloured squares show effect magnitudes; darker shades indicate higher statistical significance. Empty cells (no square) mark effects that are statistically insignificant.

For patent citations we measure patent-to-paper as a measure for translational diffusion and patent-to-patent citations as a proxy for patent quality. We find a strong positive association for the likelihood of papers building on all frontier developments, including AlphaFold 2, to be cited by patents compared to the structural biology baseline. For AlphaFold 2 and other AI frontier methods, this positive association holds for researchers and laboratories. For patent-to-patent quality, we observe that patents building on publications linked to AlphaFold 2 are almost twice as likely to be cited by other patents than baseline structural biology, with a smaller, but also positive association for other AI frontier work. However, no significant impacts are observed for researchers or laboratories who produce work linked to any frontier development and are cited in patents.

Together these results suggest several trends:

- A substantial proportion of all structural biology research results in translational outputs across clinical trials, disease related research and commercial patents. This holds true for frontier developments, including AlphaFold 2, of which a few percent of publications provide supporting inputs. This is significant as it demonstrates that AlphaFold 2 has diffused into application in a timeframe that is comparable with other frontier developments that might be more familiar to downstream researchers and innovators.
- While publications which cite frontier research appear to consistently be positively associated with translational research outputs than the structural biology baseline, this effect does not consistently translate to the researchers and laboratories producing that body of work across all measures. This is perhaps due to the small numbers or concentrated nature of applied research, meaning that systematic effects are not present or hard to detect.

4. Conclusions

4.1 Summary

Our study has examined the impact of AlphaFold 2 on experimental structural biology, academic outputs and translational results, revealing its emerging impacts on the acceleration of research and scientific impact across a diversity of fields, as compared to other structural biology developments. AlphaFold 2's impact over the time frame covered by this investigation is characterised by:

- A substantial influence on structural biology as characterised by a continued strong growth in AlphaFold 2's diffusion among the field at large. This suggests the emergence of a new platform for structural biology research alongside more established techniques implicated in other frontier developments.
- A significant and consistent boost in experimental structure submissions and a diversification in the range of proteins explored. AlphaFold 2 seems to be making a unique contribution by comparison to other frontier techniques which for example are linked to less novel protein structures. Our findings are consistent across frontier developments and for both researchers and laboratories.
- A modest increase in the academic productivity of its adopters accompanied by a growth in the citations that this research receives. The effects of counterfactual frontier techniques are similar in some cases, but AlphaFold 2's impact profile tends to be slightly stronger, and more consistent than other contemporary AI tools, particularly when field weighted citations are accounted for.
- A varied and nuanced impact on translational outputs. Papers citing AlphaFold 2 are not linked to an increase in disease-related research, but there is a positive association for researchers and laboratories building on AlphaFold 2. For clinical article citations this reverses - AlphaFold 2 associated with gains for papers, but not for researchers and laboratories. Papers, researchers and laboratories building on AlphaFold 2 are all positively associated with higher patent citation rates. Similarly mixed results appear for other frontier developments, suggesting the occurrence of translational impacts is dictated by underlying mechanisms.

Our analysis is carried out at a high level, investigating systemic impacts of AlphaFold 2 and other frontier developments in structural biology. It is not supported by additional nuanced analysis and substantial qualitative research that would be required to understand the underlying dynamics and drivers that lead to particular results and the differences between them.

4.2 Limitations

The findings presented in this report should be interpreted in light of several data and methodological limitations.

Endogeneity and Causality: Despite using robust matching methods, the study cannot fully rule out unobserved factors, such as resources available or strategic priorities, that influence researcher adoption decisions. Strict causal links cannot be established in the absence of an experimental design.

Temporal and Metric Bias: Comparing frontier technologies is complicated by varying release dates. Older methods have had significantly more time to accrue citations than AlphaFold 2. Downstream impacts, for example measured by patent data, likely underestimate applied impact by missing proprietary or unpublished applications.

Overestimation of Usage Intensity: The methodology assumes that past citations imply future adoption or influence, which likely overestimates the actual intensity of day-to-day usage. This metric should be interpreted as a signal of technological exposure rather than a measure of consistent reliance upon a category of frontier development.

Lack of Interaction Effects: The model treats the adoption of different frontier technologies as independent variables, despite the fact that researchers may use multiple advanced methods simultaneously. Consequently, the analysis cannot isolate potential synergistic benefits or substitution effects between AlphaFold 2 and other tools.

Data Constraints: Data noise regarding authorship and affiliations in OpenAlex imposes constraints on the granularity of the analysis. Additionally, limited citation intent data results in a loss of statistical power due to sample attrition, constraining our analysis.

Aggregate Nature of Findings: The focus on average effects may mask important distributional impacts and fails to explain the underlying mechanisms driving the results. As a result, while the findings highlight important dynamics, the specific interpretations and conclusions remain broadly speculative.

4.3 Discussion

This study supports the notion of high performance AI as a valuable research tool, and hints at it providing a new method of invention. We conclude this report by discussing the implications of our findings through the lens of scientific exploration, research productivity, and invention.

Scientific exploration

Our work supports the idea of AI as a tool for enhancing scientific search and discovery. The consistent association of AlphaFold 2 with more novel protein structures suggests that the tool is used to characterise proteins in less chartered parts of the protein universe. These findings suggest the capacity of AI to open up additional avenues for research within a problem space, and is in line with existing work on the use of AlphaFold 2 (Yu, 2024).

We speculate that AlphaFold 2 lowers the cost barrier and reduces the risks associated with research portfolio diversification. This is enabled by the tool being free and open to

use, and by the release of a large volume of predicted protein structures in the AlphaFold Protein Database, with self-assessed accuracy evaluations. The patterns we observe align with the notion that predictive AI tools can facilitate efficient resource allocation. By screening large knowledge spaces computationally they could allow researchers to direct experimental and applied efforts toward high value targets, avoiding diminishing returns of further exploring better known structure types (Agrawal, McHale and Oettl, 2024).

Our findings suggest that AI is a tool that can potentially mitigate the 'streetlight effect', in which it is much easier for scientists to explore areas where it is possible to build on data about existing structures (Tranchero *et al.*, 2022; Tranchero, 2024). AlphaFold 2 serves as the solution to an engineering problem that previously restricted our view of the protein population. The ability of the tool to map new spaces, and direct research efforts towards them, points more toward AI being a tool that can open up new questions. Thus, while this type of AI does not generate new scientific knowledge, as an engineering solution it can qualitatively shift the directionality of research.

It is important to note that our findings does not allay another concern that the uptake of AI will lead to scientists overly focusing on problems that are suited to AI or that overreliance on AI will diminish the ability of science to understand mechanisms that drive phenomena observed in the lab (Messeri and Crockett, 2024). While it is possible that AI-enhanced exploration may itself open up surprising new research avenues, the policy decisions and funders can also help to guard against monocultures at the macro level, by encouraging the pursuit of novel, path breaking research that might lean less heavily on AI.

Research Productivity

As well as encouraging exploration leading to the experimental discovery of new protein structures, AlphaFold 2's ability to provide highly accurate protein structure predictions has enabled researchers to produce publications of enhanced quality in marginally higher volumes. This is evidenced by a consistent enhancement in citation impact across disciplines and complemented by the substantially higher volumes of associated Protein Databank structures. From our descriptive analysis we see that it is likely that these positive effects also spill over into other fields. At a systemic level, we do not observe a factory-like churn of publications linked to AlphaFold 2. The publication rates for researchers and laboratories leveraging AlphaFold 2 increase in line with other frontier developments.

These combined aspects of AlphaFold 2's success support the 'Hassabis hypothesis', which suggests that AI-based predictive models can significantly advance scientific discovery when applied to problems with a defined objective, vast search spaces, and ample data or simulation capabilities, given the relative ineffectiveness of other modelling methods in that domain (Agrawal, McHale and Oettl, 2024). The complex and unsolved challenge of protein folding prediction satisfies these criteria, and our analysis provides evidence of the impact AlphaFold 2 has had on that problem.

Other observed increases in citation impact, such as those in clinical research, indicate AlphaFold 2's ability to not only advance fundamental research, but also to assist the wider development of impactful, pointing to spillover effects that advance science as a whole.

One nuance within these findings is that the productivity benefits of using AlphaFold 2 are typically higher for laboratories than established researchers, but the inverse is true for citation impact when measured by FWCI. Indeed, this is broadly true of all frontier developments included in the study. This suggests that laboratories are able to leverage economies of scale and flexibility in order to adopt new methods and incorporate them into larger scientific processes, which is an important aspect in the diffusion of AI methods. Other factors that may play a role include the experience contained within labs, access to knowledge through larger collaboration networks, and resources and infrastructure, which may make it easier to leverage AI methods. Nonetheless, this does not totally discount the ability of researchers who are not lab leads to leverage AI in potentially more agile ways, and produce research of high impact in their respective fields.

AlphaFold 2 has made an impact in structural biology research, including experimentation, but does not excel in all areas of impact in comparison to the counterfactual frontier developments we have chosen. Our weaker and contradictory evidence around translational impacts could be linked to lags in the translation of novel research into innovations, and the limited timeframe available since AlphaFold 2's release. Our results are indicative of the direction of travel for AI in science, but do not comprehensively cover the full breadth of impacts, nor the longer term potential benefits or drawbacks of its wider adoption. New improvements to AlphaFold 2 are enhancing its performance and addressing its limitations, potentially increasing its impact, and presenting future opportunities for study (Abramson *et al.*, 2024). Other recent developments in the ability to develop large language models with greatly reduced inputs suggest nonlinear progress in AI development may reduce some of the barriers required for researchers to build new AI tools and see them proliferate within other fields (DeepSeek-AI *et al.*, 2025).

AI Integration

Results in this study demonstrate that AlphaFold 2 is built upon by researchers to produce increased levels of experimentally verified protein structures. This finding in particular is consistent with the idea that AlphaFold 2 is a complement to experimental and domain-specific work rather than a substitute. This adds texture to the "oil and water" phenomenon by suggesting that the lack of integration between AI and existing research within a discipline is not a fixed property, but rather an attribute of specific AI, the problems they seek to solve and the task and collaboration configurations required (Duede *et al.*, 2024).

This result is complemented by the use of AlphaFold 2 across disciplines. This is in part driven by its generalisable transformer neural network architecture, which distinguishes it from other protein structure prediction innovations in structural biology. Some of our

findings point to impacts where other AI innovations have struggled to make inroads, that in some cases it matches the impact of frontier non-AI developments in structural biology. The fact that this impact also stretches beyond biochemistry suggests that AlphaFold 2 presents a significant advancement on the state of AI within the field, while also pushing the state-of-the-art for AI in general. This highlights the potential compounding effect of AI advancements across science if adopted as a general method of invention.

Finally, we see emerging signals of AlphaFold 2's influence in more applied research and development. The somewhat contradictory results, relatively small numbers of adjacent translational outputs, and the fact that other frontier developments prove similarly likely to contribute to applied impacts suggests a number of dynamics. The first is that there is an obvious lag time between new transformational innovations and effects on applied and commercial innovation activities. We speculate that despite being a technology of transformative potential, AlphaFold 2 is experiencing this lag in part because the research ecosystem will need to undergo certain reconfigurations to realise its full potential. Laboratories and the individuals within them will likely need to engage in new specialisms, develop new collaboration patterns, and identify ways to manage the end-to-end processes that employ new AI tools to create applied knowledge and innovative outputs.

Supporting Science with AI

Based on this investigation and our findings, we suggest three priority actions that might enhance the impacts of AI in science, based on our findings relating to AlphaFold 2:

- Ensure that funds to support fundamental scientific discovery are balanced with the desire to develop and capitalise on new AI tools. Funders should be aware of the benefits and limitations of AI and design calls for proposals to promote distinct research, as well as encourage better integration. In this way the science ecosystem can efficiently support the translation of newly mapped problem spaces into new knowledge.
- For successful AI prediction tools that are sufficiently accurate and demonstrate strong adoption among researchers, early efforts should be made to ensure a proportion of AI R&D is focused towards creating infrastructure and open access tools required for researchers to take advantage of them. Our findings suggest that established researchers can use AI tools to take research fields in new and applied directions, which can be further supported by lowering the barrier to entry.
- Where AI leads to an explosion of the mapped space in a field, new tools for knowledge management and exploring that space should be developed. For proteins there are structure databases, but linking these with interactive tools that link structures and academic knowledge in interactive ways, using natural language queries and building on existing research workflows could be powerful.

We also suggest three additional areas for research that would build on our work here and fill important evidence gaps. These are: economic analysis to understand how AI impacts

the allocation and efficiency of research spending; qualitative research to more deeply understand how scientists are integrating AI into their workflows; and collaboration analysis using co-authorship data to understand how team composition and multidisciplinary integration are being affected at the ecosystem scale.

Bibliography

Abadi, M. *et al.* (2016) 'Tensorflow: A system for large-scale machine learning', in *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, pp. 265–283.

Abramson, J. *et al.* (2024) 'Accurate structure prediction of biomolecular interactions with AlphaFold 3', *Nature*, 630(8016), pp. 493–500. Available at: https://doi.org/10.1038/s41586-024-07487-w.

Agrawal, A., McHale, J. and Oettl, A. (2018) 'Finding Needles in Haystacks: Artificial Intelligence and Recombinant Growth', *NBER* [Preprint].

Agrawal, A., McHale, J. and Oettl, A. (2024) 'Artificial intelligence and scientific discovery: a model of prioritized search', *Research Policy*, 53(5), p. 104989. Available at: https://doi.org/10.1016/j.respol.2024.104989.

AlQuraishi, M. (2020) 'AlphaFold 2 @ CASP14: "It feels like one's child has left home."', *Some Thoughts on a Mysterious Universe*, 8 December. Available at: https://moalquraishi.wordpress.com/2020/12/08/AlphaFold 2-casp14-it-feels-like-ones-child-has-left-home/ (Accessed: 28 September 2023).

Andrychowicz, M. *et al.* (2023) 'Deep Learning for Day Forecasts from Sparse Observations'. arXiv. Available at: https://doi.org/10.48550/arXiv.2306.06079.

Bail, C.A. (2024) 'Can Generative AI improve social science?', *Proceedings of the National Academy of Sciences*, 121(21), p. e2314021121. Available at: https://doi.org/10.1073/pnas.2314021121.

Bertoline, L.M.F. *et al.* (2023) 'Before and after AlphaFold 2: An overview of protein structure prediction', *Frontiers in Bioinformatics*, 3. Available at: https://doi.org/10.3389/fbinf.2023.1120370.

Bianchini, S., Müller, M. and Pelletier, P. (2022) 'Artificial intelligence in science: An emerging general method of invention', *Research Policy*, 51(10), p. 104604. Available at: https://doi.org/10.1016/j.respol.2022.104604.

Bianchini, S., Müller, M. and Pelletier, P. (2024) 'Drivers and Barriers of AI Adoption and Use in Scientific Research'. arXiv. Available at: https://doi.org/10.48550/arXiv.2312.09843.

Brynjolfsson, E., Li, D. and Raymond, L. (2023) 'Generative AI at Work'. arXiv. Available at: https://doi.org/10.48550/arXiv.2304.11771.

Brynjolfsson, E. and Mitchell, T. (2017) 'What can machine learning do? Workforce implications', *Science*, 358(6370), pp. 1530–1534.

Brzezinski, P. *et al.* (2024) *The Nobel Prize in Chemistry 2024*, *NobelPrize.org*. Available at: https://www.nobelprize.org/prizes/chemistry/2024/popular-information/ (Accessed: 22 November 2024).

Callaway, E. (2020) '"It will change everything": DeepMind's AI makes gigantic leap in solving protein structures', *Nature*, 588(7837), pp. 203–204. Available at: https://doi.org/10.1038/d41586-020-03348-4.

Chenthamarakshan, V. *et al.* (2023) 'Accelerating drug target inhibitor discovery with a deep generative foundation model', *Science Advances*, 9(25), p. eadg7865. Available at: https://doi.org/10.1126/sciadv.adg7865.

DeepSeek-AI *et al.* (2025) 'DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning'. arXiv. Available at: https://doi.org/10.48550/arXiv.2501.12948.

Duede, E. *et al.* (2024) 'Oil & Water? Diffusion of AI Within and Across Scientific Fields'. arXiv. Available at: http://arxiv.org/abs/2405.15828 (Accessed: 4 June 2024).

Duran-Frigola, M., Mosca, R. and Aloy, P. (2013) 'Structural Systems Pharmacology: The Role of 3D Structures in Next-Generation Drug Development', *Chemistry & Biology*, 20(5), pp. 674–684. Available at: https://doi.org/10.1016/j.chembiol.2013.03.004.

Elofsson, A. (2023) 'Progress at protein structure prediction, as seen in CASP15', *Current Opinion in Structural Biology*, 80, p. 102594. Available at: https://doi.org/10.1016/j.sbi.2023.102594.

Eloundou, T. *et al.* (2023) *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*, *arXiv.org*. Available at: https://arxiv.org/abs/2303.10130v5 (Accessed: 26 September 2024).

Evans, R. *et al.* (2022) 'Protein complex prediction with AlphaFold-Multimer'. bioRxiv, p. 2021.10.04.463034. Available at: https://doi.org/10.1101/2021.10.04.463034.

Gao, J. and Wang, D. (2024) 'Quantifying the Benefit of Artificial Intelligence for Scientific Research'. arXiv. Available at: https://doi.org/10.48550/arXiv.2304.10578.

Gargiulo, F. *et al.* (2022) 'A meso-scale cartography of the AI ecosystem'. arXiv. Available at: http://arxiv.org/abs/2212.12263 (Accessed: 19 September 2024).

Hammond, M. (2023) 'The path to applied science co-pilots.', *Accelerating Applied Science by Deep Science Ventures*, 15 December. Available at: https://acceleratingappliedscience.substack.com/p/the-path-to-applied-science-co-pilots (Accessed: 26 November 2024).

Hazra, S. and Patra, S. (2021) 'Importance of protein structure and function in pathogenesis: Highlights on the multifaceted organism *Mycobacterium tuberculosis*', in S. Das and H.R. Dash (eds) *Microbial and Natural Macromolecules*. Academic Press, pp. 775–809. Available at: https://doi.org/10.1016/B978-0-12-820084-1.00030-2.

Higgins, M.K. (2021) 'Can We AlphaFold Our Way Out of the Next Pandemic?', *Journal of Molecular Biology*, 433(20), p. 167093. Available at: https://doi.org/10.1016/j.jmb.2021.167093.

Hill, R. and Stein, C. (2020) 'Race to the Bottom: Competition and Quality in Science'.

Jakubik, J. *et al.* (2023) 'Foundation Models for Generalist Geospatial Artificial Intelligence'. arXiv. Available at: https://doi.org/10.48550/arXiv.2310.18660.

Jumper, J. *et al.* (2021) 'Highly accurate protein structure prediction with AlphaFold', *Nature*, 596(7873), pp. 583–589. Available at: https://doi.org/10.1038/s41586-021-03819-2.

Krenn, M. *et al.* (2022) 'On scientific understanding with artificial intelligence', *Nature Reviews Physics*, 4(12), pp. 761–769. Available at: https://doi.org/10.1038/s42254-022-00518-3.

Merchant, A. *et al.* (2023) 'Scaling deep learning for materials discovery', *Nature*, 624(7990), pp. 80–85. Available at: https://doi.org/10.1038/s41586-023-06735-9.

Messeri, L. and Crockett, M.J. (2024) 'Artificial intelligence and illusions of understanding in scientific research', *Nature*, 627(8002), pp. 49–58. Available at: https://doi.org/10.1038/s41586-024-07146-0.

Paszke, A. *et al.* (2019) 'PyTorch: An imperative style, high-performance deep learning library', in *Advances in neural information processing systems 32*. Curran Associates, Inc., pp. 8024–8035. Available at: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

Pedregosa, F. *et al.* (2011) 'Scikit-learn: Machine learning in python', *Journal of machine learning research*, 12(Oct), pp. 2825–2830.

Perrakis, A. and Sixma, T.K. (2021) 'AI revolutions in biology', *EMBO reports*, 22(11), p. e54046. Available at: https://doi.org/10.15252/embr.202154046.

Saplakoglu, Y. (2024) *How AI Revolutionized Protein Science, but Didn't End It*, *Quanta Magazine*. Available at: https://www.quantamagazine.org/how-ai-revolutionized-protein-science-but-didnt-end-it-20240626/ (Accessed: 17 October 2024).

Savage, T. *et al.* (2024) 'Machine learning-assisted discovery of flow reactor designs', *Nature Chemical Engineering*, 1(8), pp. 522–531. Available at: https://doi.org/10.1038/s44286-024-00099-1.

Schweidtmann, A.M. *et al.* (2018) 'Machine learning meets continuous flow chemistry: Automated optimization towards the Pareto front of multiple objectives', *Chemical Engineering Journal*, 352, pp. 277–282. Available at: https://doi.org/10.1016/j.cej.2018.07.031.

Sønderby, C.K. *et al.* (2020) 'MetNet: A Neural Weather Model for Precipitation Forecasting'. arXiv. Available at: http://arxiv.org/abs/2003.12140 (Accessed: 19 September 2024).

Sumner, J. (2024) 'The AI revolution is always just out of reach', *Nature Reviews Physics*, 6(11), pp. 1–2. Available at: https://doi.org/10.1038/s42254-024-00777-2.

Tranchero, M. *et al.* (2022) 'The Streetlight Effect in Data-Driven Exploration'. Available at: https://www.abhishekn.com/publications-all/the-streetlight-effect-in-data-driven-exploration (Accessed: 28 January 2025).

Tranchero, M. (2024) 'Finding Diamonds in the Rough: Data-Driven Opportunities and Pharmaceutical Innovation', *Academy of Management Proceedings*, 2024(1), p. 13751. Available at: https://doi.org/10.5465/AMPROC.2024.241bp.

Varadi, M. *et al.* (2022) 'AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models', *Nucleic Acids Research*, 50(D1), pp. D439–D444. Available at: https://doi.org/10.1093/nar/gkab1061.

Vaswani, A. *et al.* (2023) 'Attention Is All You Need'. arXiv. Available at: http://arxiv.org/abs/1706.03762 (Accessed: 24 November 2024).

Wang, H. *et al.* (2023) 'Scientific discovery in the age of artificial intelligence', *Nature*, 620(7972), pp. 47–60. Available at: https://doi.org/10.1038/s41586-023-06221-2.

Yu, Z. (2024) 'The Impacts of Prediction AI on Scientists: Evidence from AlphaFold'. Rochester, NY. Available at: https://doi.org/10.2139/ssrn.4711334.

Acknowledgements

The Innovation Growth Lab acknowledges support from Google DeepMind for enabling the development of this work.

We also thank Oleg Kovalevskiy, Julian Jacobs, Zanna Iscenko, and Juan Mateos-Garcia for their comments on this work.

Disclaimer

No experiments or analyses were carried out by or at Google DeepMind. All points of view and conclusions expressed are those of the Innovation Growth Lab and do not necessarily reflect the position or endorsement of Google DeepMind. Control of the methodology, research and editorial decisions were retained by the Innovation Growth Lab.

Attribution

Code

The code for this analysis can be found in the GitHub project repository.

About the Innovation Growth Lab

The Innovation Growth Lab (IGL) is a global policy lab that helps governments develop more effective policies to increase innovation and productivity. IGL's mission is to foster productive, sustainable, and inclusive economies through novel policy ideas, experimentation, data and evidence. IGL works with policymakers, researchers, practitioners and funders to address key policy challenges in the fields of science, innovation, entrepreneurship and business policies.

IGL's Data and Technology Unit supports the organisation's mission through the development of original data driven insights, the creation of tools for analysis and decision making, and by supporting policymakers and practitioners to build capabilities around the use of data.

To find out more about our work, visit innovationgrowthlab.org