

ORIGINAL RESEARCH

Widening or Deepening? How Firms Integrate AI and Large Language Models into Research and Discovery

Nishka Nerkar

North Carolina School of Science and Mathematics, 106 Eagle Meadow Court, Cary, NC, 27519, United States

Accepted: October 5, 2026

Abstract

Multiple firms spanning expertise from material science and computational chemistry to pharmaceuticals have started investing into Artificial Intelligence (AI) and Large Language Models (LLMs) to streamline and expedite their Research and Discovery (R&D) process; however, the effects of this newfound integration of AI into R&D remain unclear as it is yet to be understood whether it is changing how the process is conducted versus simply accelerating the process. This study utilizes qualitative data collected from seven semi-structured interviews with interviewees with expertise spanning from corporate research and startups to innovation economics in order to analyze the central question: does the integration of AI in R&D deepen research in already explored areas by accelerating it or does it widen the directions of research and introduce novel directions? Across all interviews, the most common limitation noted in AI's integration in R&D was human validation. Regarding the central question, the results skewed towards the deepening trend. Many interviewees mentioned that the current usage of AI causes its integration to only accelerate and improve the discovery within research directions already explored; however, interviewees explained that this was a trend based on the methodology of R&D used by most firms rather than a technological barrier of the tool itself. These results suggest that firms and policymakers should thoroughly evaluate the productivity increase of AI rather than its potential and focus on solving the several bottlenecks still not tackled by AI: validation, human judgement, and physical experimentation.

Keywords: Artificial Intelligence; Frontier Innovation; Research and Discovery; Discovery Pipeline; Qualitative Research; Research Bottleneck

1. Introduction

Frontier innovation and scientific discovery, the search for and discovery of new drugs, molecules, materials, and other advancements, is a costly and arduous process. It takes years for a singular novelty to break through, and the failure rate for these discoveries remains incredibly high. Exorbitant development outlays adds to the difficulty of the Research and Development (R&D) process. Recently, Artificial Intelligence (AI), specifically Large Language Models, have been claimed to be a solution to these challenges. Artificial Intelligence is believed to speed up the challenging trial-and-error process of R&D from years to a few days, introduce novel research ideas that haven't been explored yet, steer R&D into unknown and unexplored directions, and substantially decrease the cost barrier of R&D. However, it has yet to be

discovered how the implementation and use of Artificial Intelligence has changed how R&D firms and companies execute and adapt to this new technological advancement of AI.

This paper explores a specific facet of this question: is AI deepening the research being conducted by expediting the R&D process and furthering research within the directions already chosen and pursued or is it broadening the horizons of research by introducing and pursuing new and novel directions of research that would have otherwise not been delved into? These two results heavily differ, and understanding which way AI implementation leans will provide great insight regarding how AI will mold and affect the R&D process in the long-term.

The current research regarding this topic tends to be in one of two categories. The first option is specialized focusing on how AI performs in specific tasks: protein folding accuracy, chemical property prediction, compound matching. The second option focuses on AI's effect on a macro scale by analyzing patterns within patent releases or publications to understand whether AI is quickening the R&D process within its decided fields. These categories highlight a gap within the current research which is qualitative, practitioner-level evidence on how AI's usage decisions are currently being made and what the noticeable compromises are being made on speed, accuracy, delivery, etc. This paper targets this exact gap by utilizing evidence from focused, structured interviews of practitioners utilizing, building, and studying AI's integration into firms' R&D process in a pharmaceutical, material science, and computational chemistry setting.

This analysis finds two major findings. First, all interviewees converged heavily on the concept that human assessment, specifically, the ability of the human to determine which AI-generated hypotheses to pursue or not, is the greatest limiting factor in the R&D process. Additionally, as AI improves and grows, this role of the human in the R&D process will only become more important. Secondly, regarding the central question of whether AI is deepening or widening the R&D process, the interviews generally leaned to the deepening end: depicting that AI's integration into the R&D process is accelerating and bettering research in the directions already explored rather than exploring novel and unexplored research directions. However, multiple interviewees explained that this deepening pattern is a result of the manner in which most firms choose to utilize AI rather than its limited capabilities. These explanations suggest that this current deepening tendency is subject to change as AI's integration gets adapted in more and different ways.

The rest of the paper goes as follows. Section 2 looks into related literature on the role of general purpose technology and AI in scientific discovery and research. Section 3 focuses on the study's methodology. Section 4 describes the findings of the study organized based on five prevalent cross-interview themes. Section 5 looks into the implications of the findings, and Section 6 concludes the overall study.

2. Literature Review

2.1 AI Acts as a General Purpose Technology

Many economists argue that some technological tools are renowned not for their specific specialization but rather their generality: their ability to be applied to multiple fields and change the structures of common industrial processes. Electrification, the steam engine, and the computer are all common examples of what economists refer to as general purpose technologies (Bresnahan and Trajtenberg 1995). A large amount of economic research points to AI and large prediction based machine learning tools, an example being today's LLMs, as classifications under this class of general purpose technologies (Henderson and Cockburn 1994). The main argument using this information applied to AI's integration in R&D would be that AI does not simply improve R&D but rather it alters the sequence of the R&D pipeline allowing for discoveries to be made earlier without the conformational stages present in the original R&D pipeline. This argument would imply that AI's contributions in R&D should exceed merely accelerating the R&D pipeline and, thus, truly widen the directions of research explored.

2.2 Potential Systematic Bias in AI-Generated Ideation

A different research paper related to this same study of exploration describes the process of innovation as a search for ideas across a large expanse of possibilities in which firms and labs must choose which candidate to pursue given the restraints of time and resources (L Fleming 2001). This research provides a strong and useful connection to this study: with the extent on which AI is trained by past literature, research, and data, it may systematically embed a bias for research within directions already explored suggesting ideas that ultimately accelerate research that would have most likely been explored in the near future, or, in other cases, it may also create a bias against the direction of research explored and provide candidates so far from the current direction of exploration that they become unfeasible and impractical.

2.3 Current Evidence of AI's Integration in Research and Discovery

The evidence suggesting that AI has contributed to R&D is factual and accurate; however, it may not be as substantial as common belief. A strong example of AI's contribution is the well-known protein structure prediction system named AlphaFold that won the creators a Nobel Prize in Chemistry, the first notable case in which an AI system made a Nobel-level scientific contribution. The LLM-based approach to retrosynthesis planning is another contribution made by AI, after AlphaFold, in which a large text corpora of literature and patent records is utilized in order to suggest synthetic routes to target molecules and the early foundation models of chemistry and materials science. However, this tool becomes difficult to implement at a large scale due to the computational and data requirements. Another example can be the development of the mRNA vaccine during the Covid-19 pandemic in which the computational tools utilized shortened the discovery pipeline by a measure of a few weeks, which although was not a substantial shortening still remained a signal of great progress and AI-induced acceleration.

2.4 How Does This Paper Address a Gap?

This literature review reveals a large gap. Records of technical work conducted by AI provides a measure of what AI contributes in specific, benchmarked tasks. Documents focusing on the impact of AI at a macro level explain AI's contribution on the level of a large research field or an economy. Neither methods of research provide insight into how specific firms or labs choose to utilize AI as a research tool, what balances they believe they have to strike between speed and accuracy, and how they believe AI has been contributing to their research overall in terms of efficiency and direction. This paper addresses this gap by providing qualitative evidence from experts across the various points of the discovery pipeline.

3. Methodology

3.1 The Research Design

This investigation has an exploratory qualitative design through the utilization of semi-structured interviews of industry and academic experts. This design is suitable for this study given its purpose of understanding how industry professionals perceive and maneuver the implementation of AI into the R&D process. A quantitative approach to this question would have limited the amount of detail and depth of understanding this paper would achieve given the early stages of inquiry this research question stands in.

3.2 Sample Description

This study conducted seven interviews with individuals holding different positions within the utilization of AI in R&D. These seven individuals were chosen purposefully to gain insight on AI's integration with various outlooks: founders and builders for AI-for-science tools, corporate scientists, computational chemists who directly utilize these tools, and academic economists who study the innovation and R&D process at a larger level. This diverse body was intentionally chosen so that the central question of whether AI's integration in the R&D process has deepened or widened search can be analyzed through the multiple lenses of the R&D pipeline.

Table 1. Interviewee Breakdown

ID	Position	Expertise	Institution Type
P1	Research Organization Leader/Founder	Funding of Cross-Disciplinary Frontier Innovation	Research Organization, Non-Profit
P2	Advisor/Senior Computa- tional Chemist	Drug Discovery in Pharmaceuti- cals	Large Corporation
P3	Professor of Innovation Eco- nomics	Academic	University
P4	Researcher, Innovation and Hybrid Intelligence	Academic/Applied Research	Independent Research Practice
P5	Founder, AI-In-Science Com- pany	Drug Discovery, Material Sci- ence	Startup Company
P6	Graduate Researcher in Com- putational Chemistry	Computational Chemistry in Academia	University-Based Lab
P7	AI/Machine Learning Re- searcher	Clinical Pharmaceutical Devel- opment	Large Corporation

3.3 Anonymization and Confidentiality

In order to ensure interviewee confidentiality, all subjects are referred to by anonymous markers (P1–P7). Additionally, specific indicative information about the interviewees, such as company name, specific project name, and niche field contribution, has been generalized in order to avoid identification through deduction. Interview P6 was actually conducted with a substitute interview due to the request of the original interviewee. The substitute requested that the interview not be recorded for privacy reasons; therefore, the analysis of interview P6 is based on notes taken throughout the interview rather than word-by-word transcription. The P6 interview also reflects the insight of the substitute rather than the original interviewee requested.

3.4 Interview Design

All interviews utilized the same semi-structured construct to design the set of questions asked. Each interview question set contained questions pertaining to four main ideas:

1. Provide an example or experience in which the interviewee has utilized AI, and, specifically, share a time where the AI's results and usage deviated from the interviewee's expectation.
2. Explain the human's role in the processes conducted by AI, the shift in the role as AI expands and grows, and who, specifically, decides how the suggestions/ideas created by AI are utilized.
3. Explain the value of data, proprietary resources, and program when understanding which firms reap which benefits from the integration of AI.
4. Explain, in the interviewee's opinion, whether the role of AI in the R&D process is widening or simply accelerating the R&D process.

Each interviewee's specific questions were based on these concepts but tailored to their specific experiences with the utilization of AI.

3.5 Analysis Method

The transcripts and notes of each interview were analyzed and organized thematically and by main idea rather than an approach focused on the interviews individually. This analysis followed a method with three steps as listed below.

Stage 1: Coding for Individual Responses. Each transcript was read fully once followed by a re-read question-by-question. In the second re-read individual statements about AI's usage were tagged with small descriptions explaining their meaning (e.g., "AI limited to already explored areas," "AI needs a human to evaluate the suggestions," "barriers with proprietary data"). The statements were labeled based on what they stated about AI rather than what question they responded to. This is because, oftentimes, responses provided extra detail about AI's capabilities that didn't always pertain to the question's topics.

Stage 2: Grouping Common Themes. After all tags were created, they were all analyzed across the seven interviews to find which tags repeated. Tags with a common repeated message were grouped into one singular theme. For example, tags describing how scientists had to consistently catch errors by the AI model, evaluate whether a hypothesis was feasible, and reject hypotheses that overstated its impact were grouped into Theme B (the humane judgement/validation theme). This is because each tag depicted the same common pattern from different perspectives and situations. Repeated themes were only published if at least three out of the seven interviewees expressed that pattern in some way. This was to avoid explaining patterns that were biased towards the perspective of only one or two interviewees. The five themes written later all met this threshold.

Stage 3: Accounting for Contradiction. When different interviewees' perspectives contradicted each other or when a single interviewee provided opposing observations (e.g., P6 stating that the current AI usage was limited to the directions explored in the past while also explaining this pattern as a result of its implementation rather than a technological limitation), both perspectives were included and analyzed rather than trying to simplify them into a single position. Disagreement was given as much importance as a recurrent theme and recognized in the disagreement paragraph for each theme listed in Section 4.

3.6 Coding Framework

Since this research relies on the distinction between AI deepening research versus widening it, it was essential that the distinction be made clear. Thus, a coding process was used to distinguish responses to either side to allow for thorough analysis rather than loose interpretation.

A response was classified as evidence of *deepening research* if the AI or LLM tool in use was described as: improving the accuracy, speed, or efficiency of research within a direction that the interviewee team/interviewee had already started pursuing; produced hypotheses, candidates, or search results within a limited scope, group, or research direction that the interviewee regarded as familiar or had already begun exploring; or explicitly stating AI's role as accelerating research that would have been completed regardless of its integration (e.g., P6 stating that AI accelerates work that researchers "would have done anyway").

A response was classified as evidence of *widening research* if the AI or LLM tool in use was described as: ideating a candidate, hypotheses, or suggestion that would not have come up without AI's integration; eliminating previous barriers ("such as expertise with a specific tool or literature group") allowing for the research team to pursue ideas previously unavailable to them; or explicitly stating AI's impact as one that allows/enables possibilities that were out of reach in their previous research trajectory/system.

Additionally, two more coding rules were utilized for proper analysis. The first rule is that when an interviewee made a statement about the future capabilities of AI (e.g., "AI will be able to...") the statement was classified as a conditional statement. These statements were listed in Section 4.3, rather than being listed with the current utilization and capabilities of AI. The second rule surrounded contradicting statements. If an interviewee made statements supporting both the deepening pattern and the widening pattern, both were utilized in analysis rather than forcing the interviewee to be classified in one specific group. This allowed for the paper's conclusions to be taken holistically and truly representative of the current situation rather than a result of assumption and simplification.

4. Findings

4.1 AI's Productivity vs. Potential

Interviewees unanimously emphasized the gap within the discussion surrounding AI's capabilities and contribution to the R&D process and its true assistance in the process. Interviewee P3 provided a clear-cut description of the situation currently going on surrounding AI's contribution by referencing economist Robert Solow and his famous statement about computers stating that they are seen "everywhere but in the productivity statistics." P3 used this statement to depict the divergence, currently noted in the world, between the current hype surrounding AI's contribution in the R&D environment and its actual impact in the economy and research productivity. Both P1 and P5, when asked to provide a specific example

of when AI visibly contributed to the R&D process, provided the example of AlphaFold's contribution to protein structure prediction that went on to win a Nobel prize. Both P1 and P5, however, heavily emphasized that this contribution was simply a mere example of one contribution rather than a groundbreaking hallmark discovery that completely changes the value of AI. P5 discussed how the change after the shift of consumer-focused LLMs has "been more in the perception of what AI is perceived to be able to do" rather than the actual delivered results and productivity made possible through AI. P1, in the same way, described that they could not pinpoint an example in their own work where AI made a discovery by itself. They, instead, explained that the role of AI was, moreso, about easing the R&D process by improving efficiency and accelerating the process rather than coming up with groundbreaking inventions independently.

The concrete evidence of value provided by AI in R&D tended to be much smaller and confined than popular belief. P3 noted the shortened timeline for the mRNA vaccine prospect in the Covid-19 pandemic as a real but small contribution of AI-acceleration in the R&D process. During the pandemic, during a period where time was a huge factor, AI-acceleration allowed for scientists to cut two weeks off the R&D process timeline. This case depicts where speed helped contribute to the R&D process without drastically changing the methodology or way in which R&D is being conducted. This arrangement in which AI allowed for bounded acceleration in the R&D process was a concept cited multiple times throughout multiple interviews suggesting that acceleration was definitively the strongest contribution made by AI in the R&D process. This finding suggests that the current discussion surrounding AI's contribution to the R&D process should take into account the difference between the tool's potential and its successful contributions.

4.2 The Continuous Research Pipeline Bottleneck: Human Judgement

The strongest and most unanimous theme within the findings covers the role of humans in the R&D pipeline. Although all interviewees hold different positions within the R&D timeline, all interviewees agreed on the idea that the role of the human in verifying and deciding what to do with AI-generated hypotheses/ideas remains the biggest limiting factor on the productivity of AI in R&D. Many interviewees added on to the argument by noting that the role of the human in this process has only been and will only continue to gain importance as AI improves.

The senior computational chemist, P2, noted that the decisions on acting on AI-generated hypotheses is usually completely made by "the scientists who's doing the work." They continued on by explaining how their career, which has been highly involved with AI in the R&D process, never focused on learning how to utilize AI but, rather, focused on applying their expertise in chemistry and scientific research on the ideas produced by LLMs. They emphasized that most hypotheses produced by AI, oftentimes, "make no sense at all" when analyzed by trained scientists. Interviewee P5 noted a similar experience in the entrepreneurial, start-up context in which they described their role as ensuring that they catch all the hypotheses/ideas that are not practical or viable with realistic constraints. Interviewee P6 that works on agentic systems within computational chemistry provided another description of the role of the human in the R&D process integrated with AI by saying that the role also heavily consists of ensuring that the hypotheses generated by AI are genuinely novel. They noted that most AI models and LLMs have strong confidence in all ideas, and do sometimes take credit for ideas that aren't actually novel and new or stem from an accidental error in the model. P6 described that the role of the scientist in catching these mistakes is extremely important to ensure progress and growth in the R&D fields.

Interviewee P1 connected these ideas to patterns within the systematic design of AI models and LLMs. P1 stated that the role of a product manager, the person with the responsibility of making decisions about how to utilize an idea/hypotheses generated by AI and what direction to take it in, only increases in importance and value. P1 explained how this role is based on the exact fact that AI models and LLMs are taking on technical tasks that lack the expertise of "taste" and coordination, the two expertise of the human in the R&D pipeline. Interviewee P4 provided strong evidence supporting this exact idea by referencing their own research on AI and human collaboration in the research of ill-posed problems such as diagnosis and forecasting. P4 recalled that, in the research, neither complete AI automa-

tion, in which AI independently conducts the research, nor the normal “human-in-the-loop” research, in which the human simply decides on what to do with AI-generated hypotheses, worked nearly as good as a unique research methodology in which the AI-judgement and human-judgement fully collaborated from the start when conducting the research. Additionally, they explained that sometimes the “human-in-the-loop” model performed the worst because at the point in which the human had to decide how to go about the hypothesis they, oftentimes, lacked the expertise to make a strong call. They ended up making decisions that were ill-informed or simply worse than the AI-model would have done by itself.

After thoroughly analyzing holistically, it is understood that the value brought by AI is extremely dependent on the human’s ability to work strongly with the model throughout to establish a research methodology that allows for both sides to collaborate equally with each other. A simple “human-in-the-loop” model, even though it allows for human input at the end, remains totally autonomous and, thus, still lacks the proper balance needed to ensure the maximum value of AI in the R&D process.

4.3 Widening or Deepening: The Current Direction

As for the paper’s central question, the interviewees that chose to explicitly answer the question leaned to the side that AI is deepening research, accelerating the research already happening, rather than widening research into genuinely novel and unexplored directions. During P1’s interview, they noted that the usage of AI in the R&D process is making discovery “deeper but narrower”; however, they specifically emphasized that this is a belief for the current usage of AI rather than a belief on the permanent role of AI in the R&D process. P6 also noted a similar pattern in AI’s contribution while coming from a different perspective. Interviewee P6 noted that the AI systems in place in their lab accelerate research that researchers “would have done anyway” rather generating new and novel ideas; however, P6 went on to emphasize that this trend was specifically due to the way in which researchers tend to utilize AI in the R&D process at this current time, implement AI as an extra layer to an already self-sufficient R&D process, rather than any structural or computational limitation on the part of the AI.

Interviewee P3 provided a structural reason for this pattern in AI’s utilization regardless of its improvement by referencing Duke Economist Sharon Belenzon’s concept of the “validation step” as the crucial step acting as a bottleneck in the R&D process and causing the separation between AI-generated hypotheses and successful innovations. Regardless of the amount of drug proposals, chemical compounds, or new theorems proposed by AI as its capabilities improve, the process of deciding and the decision of which idea goes forward and is pursued remains a lengthy process, costly in time and resources, and largely untouched by AI acceleration. This structure provides a possibility where regardless of the speed and acceleration of AI as its abilities improve, the bottleneck of validation largely controlled and carried out by the human will still remain the strongest constraint on the AI-generated hypotheses. Leading to a slow process of validation which will ultimately control and lessen the amount of ideas, novel or familiar, that ultimately become successful candidates.

Interviewee P4 suggested a possible mechanism as an explanation for this deepening trend in research within both AI-integrated R&D fields and substantially larger scale AI-independent science with large domains of information. They called on to a highly researched trend coined the “information exploration paradox” which references a pattern found across many contexts, scientific research and music rating platforms to political media consumption, found in which the increase in information available at-hand correlates, paradoxically, with a decrease in individual exploration and strong pull towards herding behavior: behavior characterized by individuals choosing the option seemingly most popular or available than continuing to explore their options individually. When specifically applied to AI, it is seen that as the cost of obtaining information and ideas approaches zero, thus the researchers could flock towards the first AI-generated suggestion rather than continuing their individual exploration which can reinforce a bias for direction-specific/familiar AI-generated hypotheses. Further research into this phenomenon is necessary to evaluate whether this explanation holds up. However, if this pattern is found to be true in AI-integrated R&D, it provides a specific explanation, without pointing to a structural or technical limitation of the tool of AI, for the tendency of AI to deepen research within existing research directions rather than explore novel research hypotheses.

4.4 Who Gets the Advantage?

Multiple interviewees emphasized that access to high-quality, oftentimes proprietary, data was the strongest determining factor for the competitive advantage held by a firm, ultimately raising the question of whether AI-integrated R&D will democratize research capacity or concentrate it within a small sub-sector of the field. Interviewee P2 mentioned that in a recent blind prediction challenge in their field, all the winners had access to proprietary data, unavailable to the larger community, that ultimately gave them a strong competitive advantage. Interviewee P2 went on to say that data quality, not algorithm choice nor model structure, was the strongest deciding factor of model quality. P5 provided this same conclusion, after coming from a different perspective, by stating that public data sets do not provide a competitive advantage as they are simply available to everyone. Proprietary data, however, provides a competitive advantage to any firms able to access or generate it because of its scarcity.

Interviewee P2's own work, however, works to counteract this disparity by generating and releasing strong, high-quality data sets on drug metabolism and toxicity in an effort to allow small firms without their own proprietary data a chance to develop high quality, competitive models with the help of these available data sets. Interviewee P5 separately identified that talent and expertise in the integration and utilization of AI in R&D remains the biggest limiting factor in the value extracted from strong, high-quality data sets in creating an AI tool. They noted that this cost remains regardless of whether the firms choose to find and utilize an AI tool outside their firm or build it within their firm. P6 added that this dynamic, present where large firms with access to proprietary data hold a competitive advantage over small firms, predates AI's integration in R&D entirely, and, although AI's integration did not start this dynamic, it is likely that it will also not end it.

4.5 Physical Barriers

The last consistent theme throughout the interviews was the challenges posed by AI when computation and physical work must finally come together. Interviewee P1 noted this challenge in their own project funding portfolio by explaining how their central question was regarding how they could integrate AI into physical laboratory experiments that didn't follow a cyclical step-by-step procedure, as the understanding is that automation for AI tools in physical experimentation with a routine cyclical procedure remains much easier than automating for a procedure lacking routine and structure. Interviewee P5 emphasized a similar skepticism explaining that fully autonomous, self-driving labs are still extremely difficult to create granted that automation depends on consistent routine and automating for several routines to accommodate the unpredictable nature of R&D would be extremely costly and harder to sustain as opposed to hiring a person to do the same work.

Interviewee P6 provided the most direct response to this question with the perspective of someone working within a chemistry lab. They stated that even if the capabilities of their AI tools and models substantially increased overnight, their limitations in the R&D process would still remain at physical experimentation as opposed to model construct or computation: two things they believe are already currently handled very well by the agentic models at play. Throughout all three interviews there remains a consistent theme: regardless of how strong the computation or model construct of the AI tools/models gets, the choke point of the discovery pipeline will remain physical experimentation and idea validation.

5. Discussion

When looked at holistically, the findings clearly support one side of the central question: currently, based on how researchers and firms choose to utilize AI in the R&D process, AI tools are deepening research in directions already explored, by accelerating the research in these directions and increasing the efficiency, rather than widening the horizons of directions explored. This conclusion is clearly seen in the conclusions of Sections 4.2 and 4.3: interviewees utilizing all sorts of perspectives within the R&D pipeline all emphasize that the most amount of value gained from AI is when it is strongly integrated with human decision and human expertise, and most claim that the research they find AI integrated in is within the bounds of research directions already explored and chosen. The validation step bottleneck suggested

as a plausible explanation by interviewee P3 provides a possible structural reason to why this deepening trend continues on even as AI's capabilities grow and improve. Looking into all possible candidates and choosing which to continue further with remains a time-consuming and costly process that AI has yet to compress, ease, or make more efficient. Given the limitations of the process, it can be understood why it is possible that even after multiple ideas have been generated and a wide variety of candidates are proposed it doesn't necessarily translate into a wider body of successful innovations.

It should be understood, however, that this trend and analysis should be looked at as a current finding rather than a permanent finding. Several interviews, most notably P1 and P6, emphasized that this trend of deepening was due to current circumstances rather than any permanent technical difficulty with the AI models. Interviewee P1 noted that the deepening pattern is a trend of "right now" and expressed hope and belief in the future possibility of change for the utilization of AI. Interviewee P6 attributed this pattern to the method that current firms and researchers use when integrating AI into the R&D process, the method of adding AI as an extra layer to an already self-sufficient R&D process rather than fully integrating it, rather than a limiting technical factor on the models' parts. Interviewee P4's reference to the information exploration paradox provides a possible reason as to why this pattern continues even as the ability of AI improves: as the cost of information and hypotheses-generation lowers, people may tend to resort to the first most available or popular option systematically allowing for herding behavior to continue. If proved true, this behavior poses a bigger challenge to AI's integration in R&D as the solution does not lie in the improved capabilities of the models but rather in the change of the human behavior running the models.

These findings connect back to the broader literature review found in Section 2 in two ways. Firstly, it gets rid of the oversimplification of AI in R&D as a general purpose technology. Although, the AI in R&D may contain the generality found in past LLMs, the evidence depicts that the effect AI has on the direction of discovery and the overall R&D process is heavily mediated by human judgement, organizational validation, and researcher support, emphasizing that these results do not simply stem from the raw technological capabilities of the tool itself. Second, the appearance of the validation step bottleneck in the R&D pipeline emphasizes the result cited by P3 referencing Belenzon's work stating that AI's effect on innovation cannot be understood by analyzing its raw technological capabilities alone. Rather, the tight points in the R&D pipeline that predate AI, experimental validation, regulatory approval, physical experimentation, continue to limit how much of AI's generative potential actually translates into physical research productivity.

These findings allow for several conclusions. For large firms, they suggest that mere investment in AI model capability is unlikely to be sufficient enough to succeed. Rather proper investment in the human validation and judgement processes to determine which candidates to pursue is extremely important. Evidence from this study suggests that successful integrations of AI are those where proper human judgement is, not replaced, but equally merged with the capabilities of AI. Interviewee P5's concerns regarding sustained training of technological professionals working on AI-tools, brings up the value and role of policymakers. Multiple interviewees emphasized that the value brought by AI in scientific discovery is found through collaboration between the capabilities of the AI model and the expertise of the human oversight which the AI models lack. This suggests that policies that approach AI's integration in scientific R&D as well as the workplace as a substitute for human action rather than a complement risk misguiding firms away from reaping the greatest value from the AI tools they utilize. Future research on this topic should include qualitative data from interviews of various experts specializing in different areas of the R&D pipeline combined with quantitative data collected longitudinally looking into patents, publications, and product statistics to explore whether the deepening pattern analyzed in this paper persists or whether the broadening research directions come forward according to the expectations of some of the optimistic interviewees.

6. Study Limitations

This study conducted contains several limitations that should be taken into account while interpreting the findings. Firstly, this study interviewed seven participants that were all purposefully chosen rather

than randomly; thus, their perspectives and reviews cannot be interpreted as a holistic representation of common belief surrounding AI's integration in scientific R&D. These interviews were all conducted at a singular point of time within a field that is rapidly changing and growing. Results reflected in this paper may not hold true years later. This is especially true given the emphasis multiple interviewees made about how their perspective represented only the current usage of AI and not persistent trends. The data collected reflects the interviewee's self-reflected perceptions on their usage of AI and its role in scientific discovery rather than verified outcomes from their firm. Thus, the data reflects their own perceptions on their work and usage of AI rather than objective measures of AI's contribution and usage in their work. Lastly, the interviewees were willing to give an interview on their experience which could imply that they were more optimistic or strong on their view of AI's usage in R&D than the general population of researchers that utilize AI. This introduces a selection bias that could not fully be removed from this study. Additionally, since the analysis was conducted by one person, it is inevitable that some human judgement could be present in theme construction.

7. Conclusion

This paper explored the central question of whether AI's integration into the scientific R&D process led to deepening of research within already explored research directions or a widening of the directions explored. Utilizing evidence collected from the interviews of seven professionals spanning from expertise in corporate research and academic economics to AI-for-science start ups, the evidence points in the direction of deepening research at this current point. This pattern has largely been sustained by a strong bottleneck, in the scientific R&D pipeline, of candidate validation as well as the human behavioral tendency to favor the most available and popular AI-generated solutions over continued novel exploration. However, it is important to note that most interviewees emphasized that this is a current trend and expressed an expectation and hope for a shift in this balance between the method of AI usage in R&D and the amount of human integration in the process. This study provides qualitative, practitioner-level evidence to a literature body that is heavily dominated by quantitative technical analysis and macro level outcome measures. The next step would be to combine this qualitative practitioner-level data with statistical productivity measures to fully and holistically analyze the trends of AI-usage and the effect of its integration in R&D to see if it evolves and shifts as per the predictions as both the models and human behavior evolves.

Funding Sources

The author declares that there were no funding sources related to this work.

Conflicts of Interest

Individuals interviewed may have financial or professional interests in the way AI integrated in the R&D process is depicted.

Acknowledgements

The author would like to thank all the interviewees for providing their insights on this research question. Additionally, the author would like to specifically thank Professor Ashish Arora from the Duke Fuqua School of Business for providing his mentorship throughout this research process.

Works Cited

- Bresnahan, Timothy F., and M. Trajtenberg. 1995. "General Purpose Technologies 'Engines of Growth?'" *Journal of Econometrics* 65 (1): 83–108. [https://doi.org/10.1016/0304-4076\(94\)01598-T](https://doi.org/10.1016/0304-4076(94)01598-T).
- Fleming, Lee. 2001. "Recombinant Uncertainty in Technological Search." *Management Science* 47 (1): 117–132. <https://doi.org/10.1287/mnsc.47.1.117.10671>.

Henderson, Rebecca, and Iain Cockburn. 1994. "Measuring Competence? Exploring Firm Effects in Pharmaceutical Research." *Strategic Management Journal* 15 (S1): 63–84. <https://doi.org/10.1002/smj.4250150906>.

Appendix: Interview Layout

Practical Use of AI: Example-Based

- Could you list an example of where AI provided a suggestion or created an idea that was utilized in your work? Please give a thorough description of AI's contribution from the start to the end of the process.
- Where in the discovery pipeline do you see AI contributing the most amount of value? Where do you see AI receiving the most amount of hype for its contribution? Is there a distinction?

Humans or AI?

- When an AI/LLM produces a candidate, who, if anyone, in the organization has the authority to decide whether to continue pursuing the candidate or classify it as noise?
- As AI's capabilities continue to grow and improve, is the role of human validation gaining or losing importance?

Competitive Advantages? Data? Computing?

- How important is proprietary versus public data in determining the strength of an AI or LLM model?
- What is the most challenging part of integrating AI into a firm's Research and Design process? Talent? Data? Computing?

Is AI Widening or Deepening Research?

- Is AI making frontier innovation broader by suggesting new and novel ideas or is it deepening and narrowing the ideas pursued?
- What is one question about the integration of AI into Research and Design that you think about but hasn't gotten enough attention?

(The questions used to follow-up with the interviewee were tailored to the interviewee's expertise)