TOWARDS AGENTIC AI FOR SCIENCE: HYPOTHESIS GENERATION, COMPREHENSION, QUANTIFICATION, AND VALIDATION

Danai Koutra[•], Lifu Huang[§], Adithya Kulkarni[†], Temiloluwa Prioleau[*], Beatrice Soh[◇], Qingyun Wu[‡], Yujun Yan[*], Yaoqing Yang[*], Dawei Zhou[†], James Zou[$]

University of Michigan[•], UC Davis[§], Viginia Tech[†], Dartmouth College[*], National University of Singapore[◇], Penn State[‡], Stanford University[$]
Website: https://iclragenticai.github.io/
Contact: iclragenticai@googlegroups.com

1 WORKSHOP DESCRIPTION

AI models for science (Lu et al., 2024) have the potential to harness large datasets, accelerate scientific discoveries and transform numerous fields. For example, AlphaFold (Jumper et al., 2021) has accurately predicted protein structures, significantly reducing research time in biology and medicine. Similarly, AI has been used to identify compounds for diseases like COVID-19 (Ghosh et al., 2023) and cancer (Zhang et al., 2024), dramatically shortening the timelines for breakthroughs in healthcare. In mathematics, AlphaTensor (Fawzi et al., 2022) has discovered faster algorithms for matrix-matrix multiplication than Strassen's algorithm. AI systems have been applied in fields such as healthcare (Lesica et al., 2021; Ghosh et al., 2023; Zhang et al., 2024), material science (Papadimitriou et al., 2024), physics (Sulc et al., 2024), mathematics (Fawzi et al., 2022; Trinh et al., 2024), education (Paul et al., 2024), psychology (Qiu & Lan, 2024), and software development (White, 2024). However, these models are often developed in an ad-hoc manner. In contrast, agentic AI (Shojaee et al., 2024; Yang et al., 2023; Lu et al., 2024) has recently emerged as a systematic approach to advancing scientific discovery, offering the ability to autonomously perform complex tasks and interact with the researchers.

To enhance the scientific discovery process with AI, it is essential to first understand the key stages where AI-driven solutions can have a substantial impact. Scientific discovery typically involves several key steps (Hallsworth et al., 2023; Wang et al., 2023): hypothesis generation, quantification, experimentation, data analysis, and verification. Each of these stages is crucial for building a robust body of knowledge. Hypothesis generation involves formulating potential explanations or predictions, quantification expresses these ideas in measurable terms, experimentation gathers data under controlled conditions, data analysis extracts meaningful insights, and verification confirms findings through replication or additional evidence. Each step requires substantial human effort, which makes AI-driven solutions highly appealing (Messeri & Crockett, 2024; Lu et al., 2024). However, integrating AI into these processes presents several challenges Wang et al. (2023); Blanco-Gonzalez et al. (2023). In hypothesis generation, AI may struggle to propose truly novel ideas beyond what it has been trained on (Poincaré, 2022). Quantification is difficult when it comes to accurately modeling complex phenomena (Cornelio et al., 2023). Experimentation raises concerns as AI might suggest impractical experiments or overlook critical variables (Duan et al., 2022). In data analysis, the risks include overfitting, bias, and misinterpreting correlations (Chintala, 2022). Finally, verification is complicated by issues of replicability and consistency, making it harder to validate AI-driven results (Fok & Weld, 2023). Overcoming these challenges is essential for AI to effectively contribute to the advancement of scientific discovery.

Our mission is to foster interdisciplinary collaboration to develop fully autonomous AI systems, addressing challenges like benchmark datasets, human-AI collaboration, robust tools and methods for validating AI outputs, and trustworthiness. By tackling these issues, we can unlock AI's transformative potential in research. In this workshop, themed **Agentic AI for Science**, we will explore these critical topics and welcome diverse perspectives. We will focus on integrating agentic AI

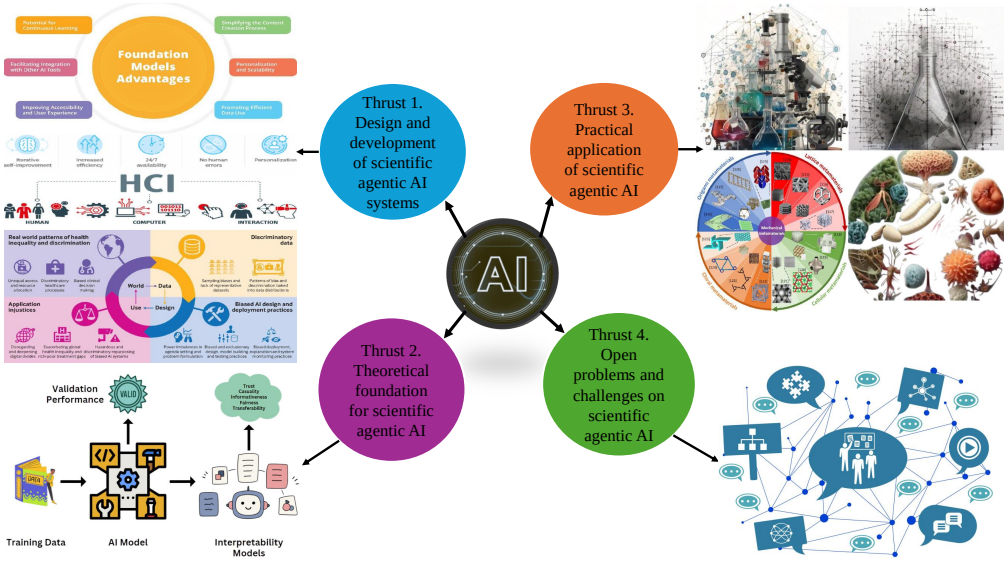

Figure 1: Illustrative figure showcasing the research thrusts addressed in the proposed workshop.

systems to enhance scientific discovery while upholding rigorous standards. For AI to contribute effectively, it must **generate novel hypotheses, comprehend their applications, quantify testing resources, and validate feasibility through well-designed experiments**. This workshop serves as a vital forum for collaboration and knowledge-sharing aimed at redefining the landscape of scientific discovery. This workshop aims to address four main research thrusts to propel future research, including (non-exclusively):

Thrust 1. Design and development of agentic AI systems for scientific discovery. The emergence of agentic AI, powered by foundation models—particularly generative models—opens up unprecedented opportunities for scientific discovery. These systems can potentially revolutionize various aspects of the scientific process, including hypothesis generation, comprehension of complex scientific phenomena, quantification, and validation. Designing and developing effective agentic AI systems for scientific discovery is both exciting and non-trivial. Pioneering work in this field has already demonstrated the promise of leveraging scientific tools, agents, and knowledge graphs. Notable examples include ChemCrow (Bran et al., 2023), which showcases the potential of AI in chemistry; Crispr-GPT (Huang et al., 2024), which applies AI to genetic engineering; and SciAgents (Ghafarollahi & Buehler, 2024), which illustrates the power of multi-agent systems in scientific discovery. These groundbreaking studies highlight the transformative potential of agentic AI in accelerating scientific progress and opening new avenues for research. Key research topics in this thrust include (but not limited to):

- Developing scientific foundation models: Tailoring general foundation models specifically for various scientific fields to enhance relevance and accuracy.
- Effective scientific tool augmentation: Enhancing existing scientific tools and methodologies with agentic AI capabilities.
- Multi-agent decomposition design: Developing frameworks for scientific hypothesis generation using multiple specialized AI agents.
- Human-in-the-loop agentic systems: Improving reliability and interpretability of AI-driven scientific discoveries through strategic human intervention.

Thrust 2. Theoretical foundation for scientific agentic AI. Developing agentic scientific AI requires methods to quantify the predictions and performance of these systems, as well as to validate the scientific hypotheses they generate. A thorough investigation of agentic scientific AI systems also demands solid theoretical foundations and tools to ensure guarantees on their behavior. To analyze and evaluate such systems, we will incorporate theoretical tools in modeling, logical rea-

soning, model validation and diagnosis, interpretable AI, and other general methods that can provide guarantees on agentic systems. Key topics in this area include, but are not limited to, the following:

- Theoretical foundation: Statistical models and theories of agentic scientific AI, such as theoretical studies on in-context learning, multi-agent communications, game theory, physics-informed hard and soft optimization constraints, and neural operators.
- Logic reasoning: Inductive, deductive, and abductive reasoning; Bayesian reasoning and probabilistic programming; neural-symbolic approaches.
- Model quantification, validation, diagnosis: Theory-driven metrics for quantifying AI system performance; self-evaluation of LLMs; data valuation and data-centric AI; diagnostics for data, architecture, and training processes; creation of standardized benchmarks for evaluating the validity of scientific hypothesis generation; scientific facts and hallucination.
- Interpretable AI: Approaches for explaining agentic AI system behaviors; quantifying trust, safety, and transparency; mechanistic interpretability.

Thrust 3. Practical application of scientific agentic AI. Deploying agentic AI systems in practical scientific research across diverse domains presents numerous challenges, particularly due to the need for domain-specific adaptation such as the unique data formats and model constraints of each scientific field. Bias in training data poses a significant risk, especially in sensitive domains like medicine. Trustworthiness and explainability are essential for scientists to confidently integrate AI-generated hypotheses and solutions into their research. Furthermore, ethical considerations arise when AI systems potentially automate research decisions that may impact public health, policy, or environmental outcomes, underscoring the importance of responsible AI deployment in science.

- Domain-specific model adaptation: Adapting agentic AI models to handle domain-specific data formats, workflows, and tools across various scientific fields; transfer learning and data-efficient fine-tuning;
- Bias detection and mitigation: Identifying and mitigating bias in training data, model design and outputs; fairness-aware AI systems for sensitive domains like healthcare and social science.
- Robustness, trustworthiness and explainability: Methods for improving the transparency and explainability of agentic AI systems in scientific research; uncertainty interpretation and quantification.
- Ethical considerations and responsible use of agentic AI in sensitive research areas; development of AI governance models to ensure accountability and human oversight in automated scientific workflows.

Thrust 4. Open problems and challenges on scientific agentic AI. Despite the promising potential of agentic AI in scientific discovery, many open problems and challenges remain to be addressed. These may include,

- Automatic curation of domain-specific scientific domains and integration of the knowledge into agentic AI systems.
- Advanced mechanisms of multi-agent collaboration in scientific discovery, with considerations of their scalability and computational efficiency.
- Continual evolution and learning of agentic AI systems; Mechanisms for updating models and improving performance based on experimental results, new data and discoveries.
- Validation and reproducibility of results generated by agentic AI systems

Addressing these challenges will require interdisciplinary collaboration between AI researchers, domain scientists, ethicists, and policymakers. Successfully overcoming these hurdles will pave the way for agentic AI to become an indispensable tool in advancing scientific knowledge and accelerating discovery across various fields. This workshop aims to bring together researchers, practitioners, and scientists from diverse fields to collectively address these challenges, foster interdisciplinary collaborations, and chart the future course of agentic AI in scientific discovery.

The proposed workshop will bring together researchers and scientists from various disciplines to enhance the development of Agentic AI systems for scientific discovery, including (non-exclusively):

- **Scientific Experts:** Experts from physics, applied mathematics, biology, chemistry, material science, and psychology contribute significantly to advancing scientific AI systems. Their knowledge is vital for addressing complex challenges and fostering interdisciplinary collaboration.
- **AI Experts:** Researchers in Natural Language Processing and Data Mining excel at processing unstructured data, enabling AI to interpret human language and extract valuable insights, which are crucial for simulating human-like reasoning and decision-making.
- **Human-AI Collaboration Experts:** Specialists from fields like human-computer interaction and communication are key to creating practical software platforms for Agentic AI systems. Their expertise ensures the development of collaboration theory and user-friendly interfaces that promote effective human-AI collaboration, making advanced technologies accessible and beneficial in real-world applications.

We believe that this workshop, by gathering researchers from diverse fields, will help identify gaps and opportunities in the existing Agentic AI research for scientific discovery.

Related Prior Symposiums and Workshops To our knowledge, there have been a few remotely related symposiums and workshops, including "Integrated Approaches to Computational Scientific Discovery" scheduled for the 2024 Fall Symposium Series[1] and the "LLMs as Research Tools" workshop held alongside CHI 2024 Aubin Le Quéré et al. (2024). Additionally, the AI4Science series has hosted several workshops, including the 1st AI4Science workshop at NeurIPS 2021, the 2nd at ICML 2022, the 3rd at NeurIPS 2022, the 4th at NeurIPS 2023, and the 5th at ICML 2024[2]. While these workshops and symposiums have made valuable contributions to the field of AI in scientific discovery, our proposed workshop distinguishes itself by focusing specifically on *agentic AI* for scientific discovery. The *agentic* nature brings in unique considerations, such as interactive and autonomous decision-making capabilities, possible integration of diverse capabilities via AI models, scientific tools, and knowledge bases, human-AI agency, and multi-agent collaboration. By exploring these unique aspects of agentic AI, our workshop aims to push the boundaries of AI-driven scientific discovery beyond the scope of previous events.

Contributions of this Workshop to ICLR Our proposed workshop promises to elevate the ICLR conference by connecting cutting-edge AI methodologies with critical scientific challenges. By concentrating on the development and application of agentic AI systems for scientific discovery, the workshop will facilitate interdisciplinary collaboration among experts in AI, natural sciences, and engineering, fostering a dynamic exchange of innovative ideas and practical solutions. Participants will engage in meaningful discussions on novel strategies for automating scientific discovery while also tackling important technical and ethical issues. By showcasing a wide range of research contributions, the workshop will underscore AI's transformative potential in scientific inquiry, solidifying ICLR's role as a leader in the convergence of AI and scientific research. Ultimately, this workshop will not only enhance the conference's offerings but also cultivate a vibrant community committed to pushing the boundaries of AI-driven scientific exploration.

2 Support for Broader Participation

Funding Support We are pleased to announce that we have secured **$1,000 in sponsorship from the Artificial Intelligence Journal and $2,500 from Amazon** for our workshop. Additionally, we are actively pursuing funding commitments from Google DeepMind and Microsoft to provide travel awards for student authors of accepted papers. Our organizing team has a proven track record of effectively managing funds for student travel awards, backed by academic affiliations that facilitate seamless fund transfers. Moreover, we are in discussions to acquire compute and API credit vouchers from Azure, Google Cloud Platform, Cohere, and OpenAI to further reward outstanding student contributions. These efforts underscore our commitment to supporting emerging researchers and fostering a diverse and inclusive academic community.

Workshop Attendance Award for K-12 and Undergraduate Students We are happy to introduce a workshop attendance award tailored specifically for K-12 students from local schools in Singapore that support underrepresented communities, including (non-exclusively) Northlight School,

[1] http://cogsys.org/symposium/discovery-2024/

[2] https://ai4sciencecommunity.github.io/

Assumption Pathway School, and Spectra Secondary School. This award also extends to undergraduate students from underrepresented communities at local institutions (non-exclusively) such as the Singapore Institute of Technology (SIT), Ngee Ann Polytechnic, and the National University of Singapore (NUS). Our initiative aims to uplift students who face financial barriers and are passionate about AI for Science. By participating in prestigious conferences like ICLR, these students will gain invaluable exposure to research opportunities and professional networks. We have proactively engaged with local schools and communities to ensure widespread interest and participation in this vital initiative. Through these awards, we aspire to empower the next generation of researchers and innovators in Singapore, nurturing their potential to make significant contributions to the field.

3 KEYNOTE SPEAKERS

Prof. Su-In Lee **(confirmed)**: Dr. Lee holds the Paul G. Allen Professor in Computer Science position at the University of Washington (UW). In addition to her primary role, Professor Lee serves as an Adjunct Professor in Genome Sciences, Electrical and Computer Engineering, and Biomedical Informatics and Medical Education at UW. She directs the Computational Molecular Biology Program, leads the AI Core at the NIH Nathan Shock Center for Basic Biology of Aging, and is the Associate Director of the Resuscitation Engineering Science Unit at UW Medicine. Professor Lee is widely recognized for her pioneering contributions to AI and biomedicine, earning numerous accolades, including the National Science Foundation (NSF) CAREER Award, the International Society for Computational Biology (ISCB) Innovator Award, the Ho-Am Prize (often dubbed the "Korean Nobel Prize"), and designation as an American Cancer Society Research Scholar. She is also a Fellow of the American Institute for Medical and Biological Engineering and has received research funding from the National Institutes of Health (NIH), the NSF, the American Cancer Society, and the Chan Zuckerberg Initiative. Her research is dedicated to advancing AI integration with biomedicine, tackling cutting-edge scientific challenges driven by AI breakthroughs. Her recent work spans a diverse array of topics, such as developing explainable AI models, uncovering the causes and treatments for complex diseases like cancer and Alzheimer's, and improving the reliability and accountability of clinical AI systems.

Prof. Markus J. Buehler **(confirmed)**: Dr. Buehler, McAfee Professor of Engineering at MIT, is renowned for his pioneering work in modeling, designing, and manufacturing advanced biomaterials. His research focuses on creating resilient materials with tunable properties across scales, linking chemical and structural features to their performance in extreme environments. His interdisciplinary approach spans mechanical, optical, and biological properties, as highlighted in his acclaimed book "Biomateriomics", which bridges diverse fields like materials science, music, and linguistics. With over 450 peer-reviewed publications, an H-index of 89, and around 30,000 citations, Buehler is a highly influential figure in his field. He has delivered over 400 invited talks and has received numerous prestigious awards, including the NSF CAREER Award, the DARPA Young Faculty Award, and the Presidential Early Career Award (PECASE). He has been recognized as a Highly Cited Researcher by Clarivate Analytics and received the Materials Horizons Outstanding Paper Prize in 2019. Buehler is editor-in-chief of the "Journal of the Mechanical Behavior of Biomedical Materials" and serves on various editorial boards. He has held leadership roles, such as president of the Society of Engineering Science, and chaired major conferences. Actively engaged in public outreach, he runs an annual materials research camp at MIT for local students and teaches a professional course on "Predictive Multiscale Materials Design". He was also elected chair of the 2021 MRS Fall Meeting, one of the largest conferences in materials science.

Prof. Sanmi Koyejo **(confirmed)**: Dr. Koyejo is an Assistant Professor in the Department of Computer Science at Stanford University and an adjunct Associate Professor at the University of Illinois at Urbana-Champaign. As the leader of the Stanford Trustworthy AI Research (STAIR) group, he is at the forefront of advancing trustworthy machine learning principles, particularly in applications related to neuroscience and healthcare. Dr. Koyejo earned his Ph.D. from the University of Texas at Austin and conducted postdoctoral research at Stanford University. His impressive accolades include

a Best Paper Award from the Conference on Uncertainty in Artificial Intelligence, the Skip Ellis Early Career Award, a Sloan Fellowship, a Terman Faculty Fellowship, an NSF CAREER Award, a Kavli Fellowship, an IJCAI Early Career Spotlight, and a Trainee Award from the Organization for Human Brain Mapping. In addition to his academic achievements, he collaborates with Google's Brain team and actively contributes to the field as a member of the Neural Information Processing Systems Foundation Board, the Association for Health Learning and Inference Board, and as the president of Black in AI. Dr. Koyejo's multifaceted expertise and commitment to fostering diversity in AI make him a vital voice in the ongoing dialogue about the future of AI in science and society.

Prof. Marinka Zitnik **(confirmed)**: Dr. Zitnik is an Assistant Professor in the Department of Biomedical Informatics at Harvard Medical School and an Associate Faculty member at the Kempner Institute for the Study of Natural and Artificial Intelligence. She is affiliated with the Broad Institute of MIT and Harvard and the Harvard Data Science Initiative. Dr. Zitnik's research aims to advance artificial intelligence to facilitate scientific breakthroughs and enhance personalized medical diagnosis and treatment. Her innovative algorithms and methods have significantly impacted the field, earning recognition from the academic, industry, and government sectors. They are utilized by leading biomedical institutions, including Baylor College of Medicine, Karolinska Institute, Stanford Medical School, and Massachusetts General Hospital. Dr. Zitnik has received numerous prestigious awards, including those from the International Society for Computational Biology, the Bayer Early Excellence in Science Award, the Amazon Faculty Research Award, and the Google Faculty Research Scholar Award. She is also a recipient of the Kavli Fellowship from the US National Academy of Sciences and the NSF CAREER Award. In addition to her academic work, Dr. Zitnik is an ELLIS Scholar with the European Laboratory for Learning and Intelligent Systems and serves on NASA's Space Biology Science Working Group. She co-founded the Therapeutics Data Commons and leads the AI4Science initiative. Her mentorship was recognized with the 2022 Young Mentor Award from Harvard Medical School.

Prof. Michael W. Mahoney **(confirmed)**: Dr. Mahoney is a distinguished researcher specializing in the algorithmic and statistical dimensions of large-scale data analysis, with an emphasis on machine learning. His expertise includes developing randomized matrix algorithms, advancing randomized numerical linear algebra, and creating geometric network analysis tools to extract structure from large informatics graphs. His work spans diverse fields such as genetics, astronomy, medical imaging, social network analysis, and internet data. Currently, he holds positions at the University of California, Berkeley, in the Department of Statistics, the International Computer Science Institute (ICSI), where he is Vice President and Director of the Big Data Group and leads the Machine Learning and Analytics Group at Lawrence Berkeley National Laboratory. Additionally, he serves as an Amazon Scholar and is part of the RISELab (formerly AMPLab). He serves on various national advisory committees, including the Statistical and Applied Mathematical Sciences Institute (SAMSI) and the National Research Council's Committee on the Analysis of Massive Data. He also runs the biennial MMDS Workshops on Algorithms for Modern Massive Data Sets, co-organized the Simons Institute's programs on the theoretical foundations of data science, and directs the NSF/TRIPODS-funded Foundations of Data Analysis (FODA) Institute at UC Berkeley.

3.1 Why we invited these speakers for the Workshop

We are proud to invite an exceptional lineup of keynote speakers who are leaders in their respective fields and exemplify the workshop's commitment to advancing Agentic AI for scientific discovery. **Dr. Su-In Lee**, a leader in integrating AI with biomedicine, will provide a unique perspective on how AI can tackle complex health challenges. Her groundbreaking research in explainable AI models and disease treatment aligns perfectly with our goal of using Agentic AI to drive scientific advancements. **Dr. Markus J. Buehler**, an expert in biomaterials, will demonstrate how AI can be leveraged in materials science to create advanced, resilient materials. His interdisciplinary approach and extensive publication record will enrich discussions on the transformative potential of AI across

scientific fields. **Dr. Sanmi Koyejo**, a pioneering figure in trustworthy AI, will share insights from his research at Stanford University, focusing on how reliable machine learning can transform neuroscience and healthcare. His expertise will be invaluable in addressing the ethical dimensions of AI in science. **Dr. Marinka Zitnik**, with her focus on AI in personalized medicine, will highlight the role of innovative algorithms in enhancing medical diagnoses. Her impressive background and recognition in the field will inspire attendees to explore the practical applications of AI in healthcare. Lastly, **Dr. Michael W. Mahoney**, renowned for his work in algorithmic and statistical analysis of large-scale data, brings a wealth of knowledge from diverse fields, including genetics and medical imaging. His experience in developing innovative algorithms will help bridge the gap between machine learning and various scientific domains, fostering interdisciplinary collaboration. Together, these speakers embody the workshop's mission to unite diverse voices and foster collaboration, setting the stage for a transformative dialogue on the future of Agentic AI in scientific discovery.

4 TARGETED AUDIENCE

We anticipate that our workshop will attract around 300 participants, including leading researchers and industry experts dedicated to advancing Agentic AI systems for scientific discovery. A significant challenge in this interdisciplinary landscape is that professionals often customize their methodologies for specific domains, hindering the development of universal frameworks. Our workshop aims to address this by bringing together engineers and researchers to foster collaboration and create comprehensive interdisciplinary guidelines that can guide future research in Agentic AI.

By bridging gaps across various fields, we aspire to promote partnerships that enhance innovation and practical applications. Our call for papers is specifically designed to connect researchers from diverse subfields within AI as well as those from external domains such as biochemistry, physics, finance, climate change, and social sciences. Through these efforts, we aim to cultivate a rich environment for knowledge exchange, ultimately advancing the impact of AI in scientific discovery and encouraging cross-disciplinary collaboration.

Plan to get audience to the Workshop To ensure strong attendance for our workshop, we have developed a comprehensive outreach plan that leverages both social media and traditional channels. We will promote the event on popular platforms like X (formerly Twitter), r/MachineLearning, and Facebook, as well as through established academic networks such as WikiCFP and those of our speakers and organizers. Additionally, we will engage with relevant affinity mailing lists, including LXAI, Black in AI, MusIML, and Queer in AI, to attract a diverse audience.

Our outreach will extend to Google groups, particularly the Women in Machine Learning Group, fostering inclusivity and engagement. With organizers collectively boasting around 10,000 followers on their active social media profiles, we will maximize our reach by encouraging them to share the workshop details within their networks.

Moreover, we are actively connecting with local schools and universities in Singapore to generate interest among students and educators. This targeted approach aims not only to ensure a wide-ranging participation but also to create a vibrant atmosphere for discussion and collaboration, making our workshop a significant event in the AI research community.

5 DIVERSITY COMMITMENT

Diversity of Topics Our workshop embraces a wide array of topics within the realm of Agentic AI research, reflecting the rich combination of ideas that can emerge from diverse perspectives. We believe that when we explore multiple dimensions of a subject, we unlock innovative pathways to solutions that address the complexities of scientific discovery. By inviting discussions that span from natural language processing to human-computer interaction, we empower participants to think creatively and collaboratively. This diversity of topics not only fosters innovation but also equips attendees with a holistic understanding of the field, inspiring them to contribute their unique insights and expertise.

Diversity of Organization Committee Our organizing committee is a powerhouse of talent and experience, hailing from seven prestigious institutions, including the Stanford University, University of Michigan, UC Davis, Virginia Tech, Penn State, Dartmouth College, and National University of

Singapore. Our team is internationally diverse, with members from the Nigeria, Greece, U.S., China, India, Singapore and Saudi Arabia, enriching our discussions with a global perspective covering four continents. This committee features a blend of voices at various career stages—tenured associate professors, assistant professors, postdoctoral fellows, and senior PhD students—each bringing their unique insights from disciplines such as NLP, HCI, interpretability, and model diagnosis. Together, they create an environment where diverse ideas can flourish and where the synergy of varied experiences leads to groundbreaking advancements. This collaboration exemplifies the strength found in unity, proving that together we can achieve remarkable heights.

Diversity of Speakers We are thrilled to present an impressive lineup of invited speakers from top-tier institutions, including Stanford University, UC Berkeley, the University of Washington, the Kempner Institute, and MIT. This diverse group which includes asian, white and black researchers comprises Vice Presidents, Professors, and Assistant Professors, featuring inspiring women who embody leadership and innovation in the field. Their collective expertise not only enhances our workshop but also serves as a powerful reminder of the value of diverse representation in academia. By sharing their unique journeys and insights, these speakers will motivate participants to break barriers and pursue their passions, fostering a community that champions equality and inspiration.

Diversity of Participants We are committed to cultivating a diverse participant base that includes individuals from various backgrounds and experiences. Our workshop actively seeks to engage participants from underrepresented communities and diverse disciplines, creating a vibrant exchange of ideas. By welcoming attendees from different cultural and academic backgrounds, we aim to foster a rich environment where every voice is heard and valued. Furthermore, we will host a special session for K-12 students and undergraduate students from marginalized backgrounds in Singapore, providing them with a platform to engage with AI research and connect with experts in the field. This diversity not only enhances discussions but also creates a network of future leaders and innovators in AI. Together, we will ignite curiosity, inspire collaboration, and empower each participant to make a significant impact in their fields.

Accessibility We are dedicated to ensuring that all individuals, regardless of their backgrounds or circumstances, have the opportunity to participate fully. Our outreach efforts include promoting the workshop on social media and establishing mentorship programs tailored specifically for underrepresented communities. By prioritizing accessibility, we strive to break down barriers and create a supportive space that inspires the next generation of researchers and innovators.

6 WORKSHOP SCHEDULE

Workshop logistics We propose a one-day workshop that will include a minimum of five keynote speeches, a panel discussion, and presentations of both invited and accepted long papers (8 pages) and short papers (4 pages) formatted according to ICLR guidelines. An awards ceremony will conclude the event. To assist participants in their preparation, we will share a reading list and additional resources ahead of time on our workshop website. The workshop will only need standard video conferencing tools, with no additional equipment required. Below is the schedule for the workshop:

Access In line with our commitment to fostering a diverse and inclusive environment, we will implement several strategies to ensure broad access to the workshop's content and discussions. Acknowledging the global nature of the research community and the potential travel constraints participants may face, we will record all keynote speeches and spotlight presentations for asynchronous viewing. This approach not only democratizes access to knowledge but also accommodates various time zones and geographical locations. By ensuring that the invaluable insights and discussions generated during the workshop are accessible to a wider audience, we aim to foster an inclusive community that benefits all, regardless of physical presence.

Encouraging Discussions among Attendees We have meticulously designed the workshop schedule to foster meaningful discussions and collaborations throughout the day. The event will begin with opening remarks, setting the stage for three compelling keynote speeches, each lasting 40 minutes, followed by 5 minutes for Q&A. This structure not only provides valuable insights from leading experts but also encourages attendees to engage with the content during a coffee break, where they can discuss the key takeaways and explore potential collaboration opportunities.

Table 1: Workshop Schedule

Singapore Time (GMT+8)	Event	Singapore Time (GMT+8)	Event
9:00-9:10	Opening remarks	14:00-15:30	Two Keynote Speeches (40 min and 5 min QA each)
9:10-11:25	Three Keynote Speeches (40 min and 5 min QA each)	15:30-15:45	Coffee Break
11:25-11:40	Coffee Break	15:45-16:45	Oral paper session (12 min talk + 3 min QA)
11:40-13:30	Poster Session and Mentoring of K-12 students and graduate students from underrepresented communities from Singapore	16:45-17:15	Interaction session with K-12 students and graduate students from underrepresented communities from Singapore
13:30-14:00	Panel Discussion	17:15-17:25	Closing remarks with awards

Next, we will feature a series of poster spotlights that showcase significant research contributions, paving the way for a dedicated poster session. This session will also emphasize mentoring for K-12 and graduate students from underrepresented communities in Singapore, creating an inclusive environment where emerging scholars can connect with experienced researchers. This initiative not only supports diversity but also enriches the workshop by fostering new collaborations and gathering constructive feedback. Following the poster session, we will delve deeper into the day's themes with a panel discussion, allowing participants to explore varied perspectives. Another round of keynote speeches will provide additional insights, ensuring a comprehensive understanding of the topics at hand.

In the afternoon, an oral paper session will feature selected researchers presenting their work in concise 12-minute talks, followed by a Q&A segment to encourage interaction and dialogue. This will be enhanced by an interactive session with K-12 and graduate students, further promoting mentorship and community engagement. The day will culminate in closing remarks that synthesize key insights and outline potential research directions. Overall, our thoughtfully curated schedule aims to maximize interaction, collaboration, and the inclusion of diverse voices, guaranteeing a rich and engaging experience for all participants. This commitment to fostering a collaborative environment is essential for advancing the field and nurturing the next generation of researchers.

7 Paper Submission

Call for Papers: Tentative important dates for paper submissions (anywhere on earth):

- Abstract Submission Deadline: January 31, 2025
- Paper Submission Deadline: February 03, 2025
- Review Bidding Period: February 03 - February 06, 2025
- Reviewer Deadline: March 02, 2025
- Acceptance/Rejection Notification Date: March 05, 2025

Template: The workshop will follow the ICLR 2025 template, which can be found at `https://github.com/ICLR/Master-Template/raw/master/iclr2025.zip`.

Page Limits: The workshop will strictly adhere to the ICLR 2025 page limits, requiring main texts to be between 6 pages for short papers and 10 pages for long papers. This policy will be rigorously enforced, and any paper exceeding the 10-page limit will be desk rejected. This rule applies to both initial submissions and final camera-ready versions. Importantly, references do not count towards

the page limit, allowing authors the freedom to provide comprehensive bibliographies. Authors can also include appendices beyond the references, but reviewers are not obligated to read them.

Types of papers that will be accepted: Our proposed workshop invites a diverse array of paper types, including original research, position papers, and survey articles, all aimed at advancing Agentic AI research for scientific discovery. We seek original research papers that present groundbreaking findings, innovative methodologies, or theoretical insights. Position papers should provide thought-provoking perspectives on emerging trends and challenges in the field, while survey articles are encouraged to offer comprehensive overviews of specific topics, illuminating current research landscapes and proposing future directions. All submissions must align with the workshop's theme and stimulate engaging discussions among participants, enhancing the collective knowledge of the community. Importantly, we will only accept papers that have not been accepted elsewhere, ensuring that all contributions are fresh and relevant. Additionally, authors can choose to submit either short (6 pages) or long (10 pages) papers, providing flexibility in presenting their work.

8 Paper Review

We will utilize the OpenReview[3] system to implement a rigorous conflict-of-interest policy for our workshop. Reviewers, including organizers, will not evaluate submissions from individuals who (1) have been colleagues within the same organization in the past three years; (2) have co-authored publications within the last three years; or (3) are currently affiliated with the same institution as the submitting authors. To ensure an unbiased review process, we will recruit reviewers from diverse institutions and varying levels of expertise. Importantly, only unpublished work will be accepted for workshop proceedings; any submissions already published elsewhere will be desk rejected. This approach reinforces our commitment to fostering a fair and transparent evaluation process, essential for advancing Agentic AI research in scientific discovery.

9 Program Team

9.1 Organizers

Prof. Danai Koutra [University of Michigan] (dkoutra@umich.edu) Google Scholar — Website **Bio:** Dr. Koutra is an Associate Director of the Michigan Institute for Data Science (MIDAS) and an Associate Professor in Computer Science and Engineering at the University of Michigan, where she leads the Graph Exploration and Mining at Scale (GEMS) Lab. She is also an Amazon Scholar. Her research focuses on principled, practical, and scalable methods for large-scale real networks, and her interests include graph summarization, graph representation learning, graph neural networks, knowledge graph mining, similarity and alignment, temporal graph mining, and anomaly detection. She has won an NSF CAREER award, an ARO Young Investigator award, the 2020 SIGKDD Rising Star Award, research faculty awards from Google, Amazon, Facebook and Adobe, a Precision Health Investigator award, the 2016 ACM SIGKDD Dissertation award, and an honorable mention for the SCS Doctoral Dissertation Award (CMU). She holds a patent on bipartite graph alignment, and has 8 award-winning papers in top data mining conferences. Over time, she has held a variety of service roles: She is an Associate Editor of ACM Transactions on Knowledge Discovery from Data (TKDD) and a program co-chair for ECML/PKDD 2023. She was a track co-chair for The Web Conference 2022, a co-chair of the Deep Learning Day at KDD 2022, the Secretary of the new SIAG on Data Science in 2021, and has routinely served in the organizing committees of all the major data mining conferences. She has worked at IBM, Microsoft Research, and Technicolor Research.

Prof. Lifu Huang [UC Davis] (lfuhuang@ucdavis.edu) Google Scholar — Website **Bio:** Dr. Huang is an Assistant Professor in the Computer Science Department at UC Davis. Prior to this, he served as an Assistant Professor at Virginia Tech and was a core member of the Sanghani AI Center. He earned his Ph.D. in Computer Science from the University of Illinois Urbana-Champaign and a Master's degree in Computer Science from Peking University. His research focuses on Natural Language Processing, Machine Learning, and Artificial Intelligence, particularly in developing efficient models and benchmarks to achieve human-level machine intelligence. His recent interests include the fundamentals and advancements of Large Language Models (LLMs) and Vision-Language Models

[3]https://openreview.net/group?id=ICLR.cc/2025/Conference

(VLMs), with applications across NLP and multimodal domains. His research has been recognized with an Outstanding Paper Award at ACL 2023 and Best Paper Award Honorable Mention at SIGIR 2023. He is a recipient of the Outstanding New Assistant Professor Award from the College of Engineering at Virginia Tech in 2024, NSF CAREER Award in 2023, Meta Research Award in 2022, and Amazon Research Award in 2021. He co-organized the "WISE-Supervision" workshop at AKBC 2022, the first and the second AI4Research Workshop at IJCAI'2024 and AAAI'2024, respectively.

Prof. Temiloluwa Prioleau [Dartmouth College] (tprioleau@dartmouth.edu) Google Scholar — Website **Bio:** Dr. Prioleau is an Assistant Professor in the Department of Computer Science at Dartmouth College. She is the founder and director of Augmented Health Lab, a consortium of interdisciplinary researchers working together to tackle grand challenges in healthcare. She is also an affiliate of the Center for Technology and Behavioral Health (CTBH). She is a Rice Academy Postdoctoral Fellow in the Scalable Health Labs. Her research is in digital health, wearable technology, and data science, with a focus on developing human-centered computing solutions that enable personalized health.

Prof. Beatrice Soh [National University of Singapore] (beatricesoh.nus@gmail.com) Google Scholar — Website **Bio:** Dr. Soh is an Assistant Professor with Presidential Young Professorship in the Department of Chemical and Biomolecular Engineering at National University of Singapore. She received her Ph.D. and M.S. in Chemical Engineering from the Massachusetts Institute of Technology (MIT) and a Bachelor's degree from Princeton University. Her research interests include Microfluidics, Polymer physics, Complex Fluids, and High-throughput experimentation. She is a recipient of the Ruth Lynden-Bell PhD Prize on Statistical Mechanics and Thermodynamics in 2022, a Rising Star in Chemical Engineering from Massachusetts Institute of Technology in 2020, and a Rising Star in Soft and Biological Matter from the University of Chicago in 2020.

Prof. Qingyun Wu: [Pennsylvania State University] (qingyun.wu@psu.edu) Google Scholar — Website **Bio:** Dr. Wu is an Assistant Professor at the College of Information Science and Technology (IST) at Pennsylvania State University, where she has been since fall 2021. Before this appointment, she worked as a Postdoctoral Researcher at Microsoft Research's New York City Lab from 2020 to 2021 under the guidance of Dr. John Langford. Dr. Wu obtained her Ph.D. in Computer Science from the University of Virginia in 2020 under the mentorship of Dr. Hongning Wang. Her research interests encompass machine learning and artificial intelligence, particularly Automated Machine Learning (AutoML) and Large Language Models (LLMs), exploring their transformative potential in intelligent information systems.

Prof. Yujun Yan [Dartmouth College] (yujun.yan@dartmouth.edu) Google Scholar — Website **Bio:** Dr. Yujun Yan joined Dartmouth College as an Assistant Professor in January 2023, following the completion of her Ph.D. at the University of Michigan, Ann Arbor. Her research focuses on the intersection of machine learning and network science, blending theoretical advancements with real-world applications in complex networks. Her work has been published in prestigious venues such as ICML, NeurIPS, KDD, and WWW. She is recognized for her significant contributions to heterophilous graph neural networks, with two of her foundational papers collectively cited over 1,200 times. Dr. Yan is also widely respected for her pioneering research on explainable graph neural networks, especially in their applications to neuroscience. She has actively contributed to workshop organization and served as a panelist at the Graph Learning Benchmarks Workshop and the PhD & Undergraduate Consortium at KDD 2023.

Prof. Yaoqing Yang [Dartmouth College] (yaoqing.yang@dartmouth.edu) Google Scholar — Website **Bio:** Dr. Yang is an Assistant Professor in the Department of Computer Science at Dartmouth College. He completed his postdoctoral studies at the RISE Lab at UC Berkeley and earned his Ph.D. from Carnegie Mellon University, following a B.S. from Tsinghua University in China. His research specializes in enhancing the reliability and transparency of machine learning models, centering on quantifying, validating, and diagnosing these models using statistical learning methods. His contributions to the field have garnered recognition as a best paper finalist at the International Conference on Distributed Computing Systems (ICDCS) and have been published in prestigious venues such as NeurIPS, IEEE Transactions on Information Theory, and CVPR. He was a committee member of the ICML CodML workshop, NeurIPS and ICLR area chair, and KDD and IEEE BigData Conference session chair.

Prof. James Zou [Stanford University] (jamesz@stanford.edu) Google Scholar — Website **Bio:** Dr. Zou is an Associate Professor of Biomedical Data Science and, by courtesy, of Computer Science and Electrical Engineering at Stanford University. His research is dedicated to improving the reliability, human compatibility, and statistical rigor of machine learning, particularly in applications pertaining to human health and disease. His algorithms are widely adopted across the tech and biotech industries. Dr. Zou earned his Ph.D. from Harvard University in 2014. He has held esteemed positions, including roles at Microsoft Research, as a Gates Scholar at Cambridge, and as a Simons Fellow at UC Berkeley. He joined Stanford in 2016 and is honored to be a two-time Chan-Zuckerberg Investigator and the faculty director of the Stanford Data4Health hub. Additionally, he is an active participant in the Stanford AI Lab. His research has garnered support from prestigious awards, including the Sloan Fellowship, the NSF CAREER Award, and AI grants from Google, Amazon, and Adobe.

Prof. Dawei Zhou [Virginia Tech] (zhoud@vt.edu) Google Scholar — Website **Bio:** Dr. Zhou is an Assistant Professor in the Department of Computer Science at Virginia Tech, where he also serves as the director of the Virginia Tech Learning on Graphs (VLOG) Lab and is a core faculty member at the Sanghani Center. His research focuses on open-world machine learning, with significant applications in hypothesis generation and validation, AI for finance, predictive maintenance, and healthcare. He has organized four workshops (TrustLOG at WWW'24 and CIKM'22, DLG@AAAI'21, and AI4FS@SDM'20) and played pivotal roles as a Vice Program Chair, Proceedings Chair, Local Chair, Social Media Chair, Publicity Chair, Session Chair and Senior Program Committee Member at numerous prestigious conferences in machine learning and artificial intelligence. His research endeavors have been supported by institutions including Virginia Tech, NSF, DARPA, DHS, the Commonwealth Cyber Initiative, 4VA, Deloitte, and Cisco. Dr. Zhou's contributions to the field have been recognized with several honors, including the Cisco Faculty Research Award (2023), selection for the AAAI New Faculty Highlights roster (2024), and the NSF Career Award (2024). Additionally, his innovative work on rare category detection leveraging human-AI intelligence was showcased by the Computing Research Association (CRA) at the 24th CNSF Capitol Hill Science Exhibition.

Workshop organizing experience for organizers Our organizing team boasts extensive experience in hosting impactful workshops and conferences across various domains. Each member has successfully managed events that foster collaboration and knowledge sharing among researchers, practitioners, and students. The organizing team's familiarity with diverse academic environments ensures we can effectively address the needs of different participants while promoting inclusivity. Furthermore, we have a track record of leveraging social media and professional networks to maximize outreach and participation. This experience positions us well to create a dynamic and enriching workshop focused on Agentic AI for scientific discovery, ensuring that all voices are heard and valued.

9.2 Confirmed Committee Members

We anticipate receiving around 50 submissions for our workshop, which will be reviewed rigorously to ensure high-quality feedback. With the plan for each submission to receive three reviews and an optimal workload of three submissions per Program Committee (PC) member, we have assembled a robust team of 54 PC members. This allows us to accommodate any unforeseen circumstances that might prevent a member from participating. Our committee is composed of experts from institutions around the globe, representing a wide range of specialties. This diverse expertise will enrich the review process, ensuring that all submissions are evaluated with the depth and insight they deserve.

- Adam Fisch (MIT)
- Adithya Kulkarni (VT)
- Ahana Gangopadhyay (WU in St.Louis)
- Aishwarya Balwani (Georgia Tech)
- Alexander Wettig (Princeton)
- Aman Madaan (CMU)
- Anna Hart (UIUC)
- Asher Trockman (CMU)
- Hyeonjeong Ha (KAIST)
- Ibraheem Moosa (PSU)
- James Zou (Stanford)
- Jiacheng Liu (UW)
- Jiajie Li (UBuffalo)
- Jianan Zhou (NTU Singapore)
- Jiawei Ma (CUHK)
- Jingling Li (UMD)

- Azmine Toushik Wasi (SUST)
- Barry Menglong Yao (VT)
- Benjamin L. Edelman (Harvard)
- Benjamin Newman (UCR)
- Charlie Victor Snell (UC Berkeley)
- Christina Baek (UC Berkeley)
- Collin Burns (Columbia)
- Danai Koutra (U Mich.)
- Daniel Y Fu (Stanford)
- Dawei Zhou (VT)
- Dingli Yu (Princeton)
- Erik Jones (UC Berkeley)
- Fatimah Alotaibi (VT)
- Giorgio Giannone (Amazon)
- Hanzi Xu (Temple)
- Haokun Liu (UNC)
- Haonan Duan (Univ. of Toronto)
- Hongyi Liu (SJTU)
- Hou Pong Chan (Alibaba)
- Kaiyuan Gao (HUST)
- Kexin Huang (Stanford)
- Kuan-Hao Huang (TAMU)
- Lifu Huang (UC Davis)
- Lijun Wu (ByteDance)
- Longfeng Wu (VT)
- Michael JQ Zhang (UT Austin)
- Michael Poli (Stanford)
- Mohna Chakraborty (U Mich.)
- Odhran ODonoghue (Oxford)
- Pascal Notin (Harvard Medical School)
- Qi Zeng (Meta)
- Qingyun Wu (PSU)
- Reza Abbasi-Asl (UCSF)
- Temiloluwa Prioleau (Dartmouth)
- Tong Zeng (VT)
- Yaoqing Yang (Dartmouth)
- Yujun Yan (Dartmouth)
- Yusuf Roohani (Stanford)

REFERENCES

Marianne Aubin Le Quéré, Hope Schroeder, Casey Randazzo, Jie Gao, Ziv Epstein, Simon Tangi Perrault, David Mimno, Louise Barkhuus, and Hanlin Li. Llms as research tools: Applications and evaluations in hci data work. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI EA '24, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703317. doi: 10.1145/3613905.3636301. URL https://doi.org/10.1145/3613905.3636301.

Alexandre Blanco-Gonzalez, Alfonso Cabezon, Alejandro Seco-Gonzalez, Daniel Conde-Torres, Paula Antelo-Riveiro, Angel Pineiro, and Rebeca Garcia-Fandino. The role of ai in drug discovery: challenges, opportunities, and strategies. *Pharmaceuticals*, 16(6):891, 2023.

Andres M Bran, Sam Cox, Oliver Schilter, Carlo Baldassari, Andrew D White, and Philippe Schwaller. Chemcrow: Augmenting large-language models with chemistry tools. *arXiv preprint arXiv:2304.05376*, 2023.

Sathish Kumar Chintala. Ai in public health: modelling disease spread and management strategies. *NeuroQuantology*, 20(8):10830, 2022.

Cristina Cornelio, Sanjeeb Dash, Vernon Austel, Tyler R Josephson, Joao Goncalves, Kenneth L Clarkson, Nimrod Megiddo, Bachir El Khadir, and Lior Horesh. Combining data and theory for derivable scientific discovery with ai-descartes. *Nature Communications*, 14(1):1777, 2023.

Jiafei Duan, Samson Yu, Hui Li Tan, Hongyuan Zhu, and Cheston Tan. A survey of embodied ai: From simulators to research tasks. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 6(2):230–244, 2022.

Alhussein Fawzi, Matej Balog, Aja Huang, Thomas Hubert, Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Francisco J R Ruiz, Julian Schrittwieser, Grzegorz Swirszcz, et al. Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature*, 610(7930):47–53, 2022.

Raymond Fok and Daniel S Weld. In search of verifiability: Explanations rarely enable complementary performance in ai-advised decision making. *AI Magazine*, 2023.

Alireza Ghafarollahi and Markus J Buehler. Sciagents: Automating scientific discovery through multi-agent intelligent graph reasoning. *arXiv preprint arXiv:2409.05556*, 2024.

Aritra Ghosh, Maria M Larrondo-Petrie, and Mirjana Pavlovic. Revolutionizing vaccine development for covid-19: a review of ai-based approaches. *Information*, 14(12):665, 2023.

John E Hallsworth, Zulema Udaondo, Carlos Pedrós-Alió, Juan Höfer, Kathleen C Benison, Karen G Lloyd, Radamés JB Cordero, Claudia BL de Campos, Michail M Yakimov, and Ricardo Amils. Scientific novelty beyond the experiment. *Microbial Biotechnology*, 16(6):1131–1173, 2023.

Kaixuan Huang, Yuanhao Qu, Henry Cousins, William A Johnson, Di Yin, Mihir Shah, Denny Zhou, Russ Altman, Mengdi Wang, and Le Cong. Crispr-gpt: An llm agent for automated design of gene-editing experiments. *arXiv preprint arXiv:2404.18021*, 2024.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *nature*, 596(7873):583–589, 2021.

Nicholas A Lesica, Nishchay Mehta, Joseph G Manjaly, Li Deng, Blake S Wilson, and Fan-Gang Zeng. Harnessing the power of artificial intelligence to transform hearing healthcare and research. *Nature Machine Intelligence*, 3(10):840–849, 2021.

Chris Lu, Cong Lu, Robert Tjarko Lange, Jakob Foerster, Jeff Clune, and David Ha. The ai scientist: Towards fully automated open-ended scientific discovery. *arXiv preprint arXiv:2408.06292*, 2024.

Lisa Messeri and MJ Crockett. Artificial intelligence and illusions of understanding in scientific research. *Nature*, 627(8002):49–58, 2024.

Ioannis Papadimitriou, Ilias Gialampoukidis, Stefanos Vrochidis, and Ioannis Kompatsiaris. Ai methods in materials design, discovery and manufacturing: A review. *Computational Materials Science*, 235:112793, 2024.

Aditya Paul, Chi Lok Yu, Eva Adelina Susanto, Nicholas Wai Long Lau, and Gwenyth Isobel Meadows. Agentpeertalk: Empowering students through agentic-ai-driven discernment of bullying and joking in peer interactions in schools, 2024. URL https://arxiv.org/abs/2408.01459.

Henri Poincaré. *The foundations of science: Science and hypothesis, the value of science, science and method*. DigiCat, 2022.

Huachuan Qiu and Zhenzhong Lan. Interactive agents: Simulating counselor-client psychological counseling via role-playing llm-to-llm interactions. *arXiv preprint arXiv:2408.15787*, 2024.

Parshin Shojaee, Kazem Meidani, Shashank Gupta, Amir Barati Farimani, and Chandan K Reddy. Llm-sr: Scientific equation discovery via programming with large language models. *arXiv preprint arXiv:2404.18400*, 2024.

Antonin Sulc, Thorsten Hellert, Raimund Kammering, Hayden Houscher, and Jason St. John. Towards agentic ai on particle accelerators, 2024. URL https://arxiv.org/abs/2409.06336.

Trieu H Trinh, Yuhuai Wu, Quoc V Le, He He, and Thang Luong. Solving olympiad geometry without human demonstrations. *Nature*, 625(7995):476–482, 2024.

Hanchen Wang, Tianfan Fu, Yuanqi Du, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, et al. Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972):47–60, 2023.

Jules White. Building living software systems with generative & agentic ai, 2024. URL https://arxiv.org/abs/2408.01768.

Zonglin Yang, Xinya Du, Junxian Li, Jie Zheng, Soujanya Poria, and Erik Cambria. Large language models for automated open-domain scientific hypotheses discovery. *arXiv preprint arXiv:2309.02726*, 2023.

Wan-Ying Zhang, Xiao-Li Zheng, Paolo Saul Coghi, Jun-Hui Chen, Bing-Jun Dong, and Xing-Xing Fan. Revolutionizing adjuvant development: harnessing ai for next-generation cancer vaccines. *Frontiers in Immunology*, 15:1438030, 2024.