AI and Science Gateways: A Promising Combination for Accelerating Science and Research Computing

Jeanette Sperhac
Sandra Gesing
Michael Zentner
Claire Stirm
jsperhac@ucsd.edu
sgesing@ucsd.edu
mzentner@ucsd.edu
cstirm@ucsd.edu
San Diego Supercomputer Center,
University of California, San Diego
San Diego, California, USA

Robert Quick
rquick@iu.edu
Pervasive Technology Institute,
Indiana University
Bloomington, Indiana, USA

Joe Stubbs
jstubbs@tacc.utexas.edu
Texas Advanced Computing Center,
University of Texas at Austin
Austin, Texas, USA

ABSTRACT

Recent advancements in Artificial Intelligence (AI), including generative systems like ChatGPT, have motivated a reevaluation of how AI methods can revolutionize research computing with possible enhancements in accessibility, resource allocation, and cybersecurity. Science gateways have long served as solutions for facilitating research by simplifying complexities associated with using advanced computing resources and promoting collaboration. Despite their shared domain in research computing, the AI and science gateway communities have a slight overlap compared to more significant potential.

To bridge this gap, the NSF Center of Excellence for Science Gateways (SGX3) has started a strategic initiative called a Blueprint Factory, to map the future of science gateways in an AI-driven ecosystem. Blueprint Factories are intensive endeavors spanning 18 months and aim to identify the technical capabilities required to support evolving scientific domains and computing resources.

As the initial event of the AI Blueprint Factory, a Birds-of-a-Feather (BoF) session was organized at PEARC'23. The session focused on emerging AI needs in research computing and their implications for science gateways. The session involved interactive discussions to explore novel modes of computing, data sharing considerations, and the potential for broadening access to compute resources through AI integration. The outcomes of this session provided valuable insights into key themes, opportunities, and challenges surrounding the integration of AI into science gateways. This paper details the methods used and the results obtained by this interactive session. It then provides an outlook on the next steps of the Blueprint Factory.

PEARC '24, July 21–25, 2024, Providence, RI, USA

ACM ISBN 979-8-4007-0419-2/24/07
https://doi.org/10.1145/3626203.3670562

CCS CONCEPTS

• **Computing methodologies** → *Artificial intelligence*; **Machine learning**; • **Information systems** → *Computing platforms*; **Web interfaces**; *Web services*; • **Theory of computation** → *Interactive computation.*

KEYWORDS

Science Gateways, cyberinfrastructure, high performance computing, machine learning, artificial intelligence, scientific software

ACM Reference Format:
Jeanette Sperhac, Sandra Gesing, Michael Zentner, Claire Stirm, Robert Quick, and Joe Stubbs. 2024. AI and Science Gateways: A Promising Combination for Accelerating Science and Research Computing. In *Practice and Experience in Advanced Research Computing (PEARC '24), July 21–25, 2024, Providence, RI, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3626203.3670562

1 INTRODUCTION

Artificial Intelligence (AI) has been influencing science and industry since the 1950s [8]. By 2011 the speed of GPUs increased to the point where they had a significant advantage in running deep learning algorithms [1]. The latest excitement around AI has been created by ChatGPT [5], which sees substantial discussion due to its public user interface and ready availability. AI methods are promising for identifying efficiencies and novel solutions in research computing, and reshaping many aspects of the field, such as accessibility and efficient resource allocation. Over two decades, science gateways have become part of the research computing landscape[3], addressing the need for usability of computational solutions, sharing simulations and data, and reproducible research.

While both the science gateway and AI communities are well anchored in the research computing landscape, these communities do not yet have much overlap [2]. The NSF Center of Excellence for Science Gateways (SGX3) [6] has set up a Blueprint Factory to consider how the research community using AI in their projects can benefit from science gateways, and in turn, the science gateways community can benefit from AI techniques in research computing. Blueprint Factories serve as one of the cornerstones of SGX3's initiative toward envisioning the future [4]. These are intensive

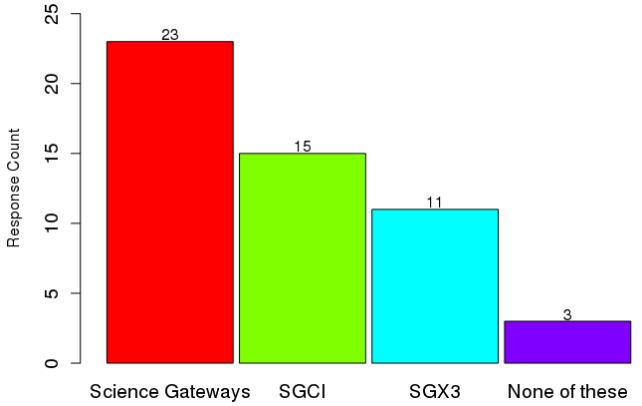
endeavors spanning approximately 18 months, focused on creating forward-looking blueprints of technical capabilities necessary for science gateways to support the future requirements of specific scientific domains or computing resources.

One part of the Blueprint Factory for AI was a Birds-of-a-Feather (BoF) session that explored the emerging needs for AI in research computing at PEARC'23 [10]. This BoF provided a forum for community exploration of emerging needs for AI-based computations and what those needs imply concerning science gateways for the next 5-10 years.

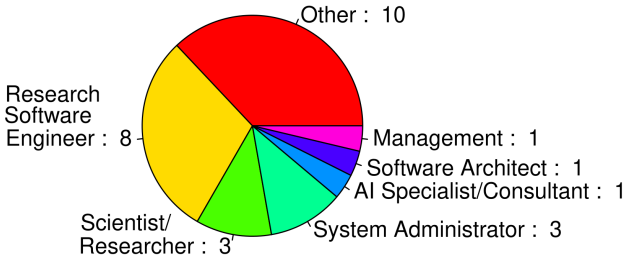
The paper will provide background about the relationship between science gateways, AI, SGX3, and its Blueprint Factories in Section 2. Section 3 will present the format of the BoF and the data collection performed during the meeting. Section 4 will summarize the analysis performed on BoF data. Section 5 presents the results. In Section 6, we offer conclusions and describe next steps.

2 BACKGROUND

Science gateways are platforms facilitating access to data, software, computing services, and specialized equipment tailored to scientific and engineering disciplines. They simplify the complexities of using research infrastructure, allowing scientists and educators to focus on their research questions and teaching. Science gateways have seen significant adoption and influence in various domains such as plant biology, phylogenetics, natural disasters, nanotechnology, chemistry, and others.

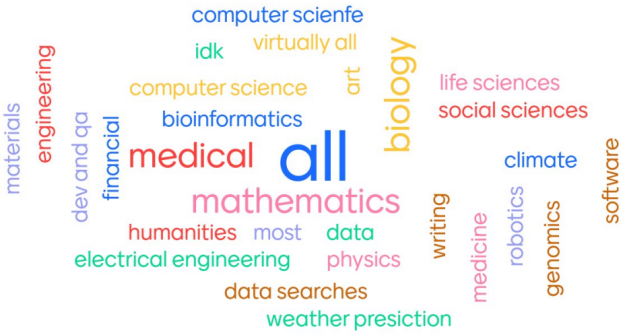


(a) Counts of BoF participants who had heard of science gateways, SGCI and/or SGX3.



(b) The self-identified primary job function of each of the 27 BoF participants.

Figure 1: PEARC'23 AI Blueprint Factory BoF participant responses.



(a) Fields that will be most impacted by AI

(b) Tasks that AI could positively influence

Figure 2: Wordclouds show BoF participant responses about the likely influence of AI on scientific fields and tasks. The size of the text corresponds to the number of responses. Each wordcloud comprises 40 responses.

2.1 SGX3

The Science Gateways Community Institute (SGCI) has supported the community in using, developing, and providing science gateways since 2016. [9] SGCI concluded its NSF-funded activities in July 2023, and continues providing paid services for clients. SGX3, a follow-on project funded in August 2022 focuses on advancing the field of science gateways with four main work thrusts:

- Serving as Community Experts, including usability, sustainability, and technology evaluation
- Growing a Diverse Community, including the Gateways annual conference
- Developing the Workforce, including hackathons, summer internships, and Focus Weeks
- Envisioning the Future, including Blueprint Factories

Lessons learned from SGCI activities informed the design of SGX3's services under these four work thrusts. Some services offered under SGCI have been retained, such as the successful Gateways conference series and Focus Week series, while others were restructured or retired.

2.2 Blueprint Factories

The concept of Blueprint Factories, unique to SGX3, forms part of the work thrust on envisioning the future. A Blueprint Factory is a concentrated activity that takes place over approximately 18 months. It aims to develop a forward-looking blueprint of technical capabilities that science gateways must offer to support future research needs.

During each Blueprint Factory, a core team will be formed consisting of five cyberinfrastructure (CI) professionals and five science domain experts. This team will gather interview contacts from their professional networks, conduct individual interviews, organize virtual focus groups, and ultimately generate a paper outlining a technological roadmap for Science Gateways over the next 5-10 years, tailored to support the domain and technologies analyzed by the Blueprint Factory.

SGX3 aims to conduct seven Blueprint Factories over five years. Their subjects will be decided from emerging topics and insights. Prospective topics include expanding the use of the ACCESS national computing resources for AI; the Materials Genome Initiative; and sustainability strategies for use by funding agencies.

2.3 Science Gateways and AI

While there is a trend of science gateways supporting AI-enabled research, offering services such as AI4Mars [7] on platforms like Zooniverse [11], science gateways still lack visibility among AI developers and users. Efforts to enhance visibility and usability within the AI community will enable the advancement of research infrastructure. Furthermore, collaborations between science gateway and AI communities are essential for driving innovation and addressing usability challenges. SGX3 addresses this topic via the Blueprint Factory on AI.

3 BOF FORMAT AND DATA COLLECTION

The PEARC'23 BoF, part of the Blueprint Factory on AI, was a highly interactive session with 27 attendees. The BoF began with an introduction to Science Gateways and conversation-starting examples of AI workflows.

The participants were asked whether they had previously heard of science gateways, SGCI, or SGX3. Except for three participants, all had heard of at least one of these concepts or projects (see Fig. 1a). Participants were then asked for their primary job function. For this question, the largest response groups were *Other*, with ten participants, and *Research Software Engineers*, with eight (see Fig. 1b). Further conversations indicated that some participants work in user support, and that an NSF director was present in the *Other* group.

Participants were then asked to predict which areas of science will be most strongly impacted by AI (see Fig. 2a), and what tasks could be most improved using the technology (see Fig. 2b). The field of science responses favored *all*, *mathematics* and *medical*. Since the majority of participants represented user support and research software engineering, their hope to improve *User Support* and *Coding* as well as *Search* reflects the backgrounds of the participants.

The majority of the BoF was dedicated to a Post-it® matrix development session. BoF participants were asked to consider four **focus questions**:

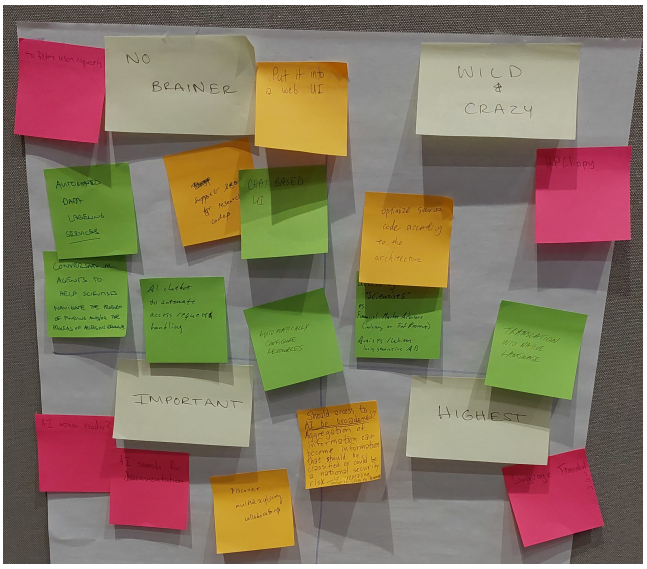


Figure 3: One of 12 idea boards collected during the BoF.

(1) What are novel modes of computing implied by AI? Which will have the greatest impact on cyberinfrastructure (CI) for gateways and national compute resources?
(2) What are the interesting aspects related to data sharing? How do the various compute loads and modes impact how data are shared, moved, and accessed?
(3) What are or should be the goals of broadening access to compute resources for AI purposes? Who can be brought into the community that previously was not?
(4) How can AI be used to make a science gateway more effective, efficient, or otherwise better?

Four groups of participants, each addressing one of the focus questions, convened for 15 minute periods during the BoF. Participants were given a short time to discuss their focus question, then to write down related ideas, each one on a separate Post-it. Participants then categorized their ideas in one of four quadrants: "No Brainer" (useful and easy to implement), "Wild & Crazy" (unconventional ideas), "Important", and "Highest Impact". Each focus area met three times with different participants, creating 12 **idea boards**, a single example of which is shown in Fig. 3.

4 DATA ANALYSIS

Participant responses during the BoF session provide a glimpse into how CI professionals think about science gateway needs and opportunities with regard to AI. We analyzed the BoF data to gain perspective on these views. Participant remarks from each of the 12 idea board sessions were collected and tabulated. A total of 314 raw remarks were collected, on average 26 remarks for each of the 12 groups. Collected remarks were summarized or generalized and organized according to themes and keywords. The top four themes that emerged from this treatment are *User Assistance*; *Usability*; *Compute and Training*; and *Data*. Section 5 details the results from

grouping participant remarks by these most frequently occurring themes and associated keywords.

5 RESULTS

Participants mentioned keywords associated with the most frequently mentioned four themes between 3 and 24 times. These themes and keywords are shown below, with the frequency of keyword mentions noted in parentheses. The results are ordered by the frequency with which they occurred, with the most mentions listed first.

Interactive User Assistance: Participants mentioned subjects centered on providing conversational, interactive assistance to gateway users 11 or more times, as follows:

- Guide data preprocessing (deduplicate, normalize, label, deidentify, classify) (24)
- Guide user selection of compute resources (scheduling, optimizing) (18)
- End-user assistant/natural language chatbot (16)
- Guide dataset selection (14)
- Guide workflow composition (14)
- Guide job formulation (11)

Gateway Usability: Introducing AI integrations for usability was mentioned 3 to 9 times, as follows:

- Cache previous computations (9)
- Natural language/query/API translation (9)
- Guide data exploration/visualization (7)
- Suggestions utilizing user history on gateway (5)
- UI customization (4)
- User support (3)

Compute and Training Concerns: Using AI to streamline compute and training was mentioned 3 to 6 times, as follows:

- Job scheduling assistance (6)
- Guide parameter selection (4)
- Predict results of simulations (4)
- Support comparison of physical models with proxy models (4)
- Guide model training (3)
- Interactive computing (3)

Data Access and Handling Concerns: AI-related data concerns and issues were mentioned 3 to 5 times, as follows:

- Data cleaning (5)
- Data ethics (5)
- Data provenance (4)
- Protected Health Information (PHI) (3)
- Access to training data (3)
- Data from edge devices (3)
- Intelligent compression (3)

Several conclusions regarding the integration of AI into science gateways can be made, based on the responses of the CI professionals who attended the BoF. Leveraging AI to provide interactive assistance with data management, resource selection, workflows, and model training are primary opportunities in science gateways. Science gateways are portals intended to democratize access to advanced computing resources, making these resources available with a lower barrier is a natural goal. Keeping democratization in mind, participants identified challenges of user assistance, computation and model training, and surfaced concerns of ethics and data provenance and handling.

6 OUTLOOK

This paper presents insights gathered from the BoF held at PEARC'23 on integrating AI into science gateways. The BoF explored emerging needs for AI-driven research and its implications for science gateways in the coming years. Through brainstorming sessions, participants addressed key focus questions, including novel modes of computing implied by AI, data sharing, broadening access to resources, and enhancing the usability of science gateways. Analysis of participant remarks revealed prominent themes and opportunities, such as task automation, workflow creation and management, model building and refinement, and dataset management. Challenges related to access to data, AI ethics, data concerns, and user support were also identified. These findings provide insights for developing AI-driven science gateways and lay the groundwork for future research and development efforts. The next phase in the AI Blueprint Factory AI includes convening the team and conducting interviews with scientists seeking to use AI techniques in their research.

ACKNOWLEDGMENTS

We would like to acknowledge NSF OAC 2231406 (SGX3).

REFERENCES

[1] Keith D. Foote. 2022. A Brief History of Deep Learning. https://www.dataversity.net/brief-history-deep-learning/

[2] Sandra Gesing, Marlon Pierce, Suresh Marru, Michael Zentner, Kathryn Huff, Shannon Bradley, Sean B. Cleveland, Steven R. Brandt, Rajiv Ramnath, Kerk Kee, Maytal Dahan, Braulio M. Villegas Martínez, Wilmer Contreras Sepulveda, and José J. Sánchez Mondragón. 2024. Science gateways and AI/ML: How can gateway concepts and solutions meet the needs in data science? In *Critical Infrastructure - Modern Approach and New Developments*. IntechOpen, London, United Kingdom. https://doi.org/10.5772/intechopen.104070

[3] Sandra Gesing, Claire Stirm, Michael Zentner, Maytal Dahan, and Linda Hayden. 2023. SGX3: Novel concepts to enhance knowledge and extend the community around Science Gateways. In *"Science Gateways 2023 Annual Conference"* (Pittsburgh, Pennsylvania). Zenodo, Geneva, Switzerland, 1–5. https://doi.org/10.5281/zenodo.10034891

[4] NSF Center of Excellence for Science Gateways (SGX3). 2023. *Blueprint Factories*. SGX3. https://sciencegateways.org/our-services/blueprint-factories

[5] OpenAI. 2022. introducing ChatGPT. https://openai.com/blog/chatgpt

[6] SGX3. 2022. Science Gateways Institute and SGX3 (SGCI/SGX3). https://sciencegateways.org

[7] R. Michael Swan, Deegan Atha, Henry A. Leopold, Matthew Gildner, Stephanie Oij, Cindy Chiu, and Masahiro Ono. 2021. AI4MARS: A Dataset for Terrain-Aware Autonomous Driving on Mars. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (Nashville, Tennessee). IEEE Computer Society, Los Alamitos, CA, USA, 1982–1991. https://doi.org/10.1109/CVPRW53098.2021.00226

[8] Alan M. Turing. 1950. Computing Machinery and Intelligence. *Mind* 49, 236 (1950), 433–460.

[9] Nancy Wilkins-Diehr, Michael Zentner, Marlon Pierce, Maytal Dahan, Katherine Lawrence, Linda Hayden, and Nayiri Mullinix. 2018. The Science Gateways Community Institute at Two Years. In *Proceedings of the Practice and Experience on Advanced Research Computing* (Pittsburgh, PA, USA) *(PEARC '18)*. XSEDE, Association for Computing Machinery, New York, NY, USA, Article 53, 8 pages. https://doi.org/10.1145/3219104.3219142

[10] Michael Zentner, Claire Stirm, Sandra Gesing, Robert Quick, and Joe Stubbs. 2023. The Impact of AI Computing Paradigms on Science Gateways and National Compute Resources. https://doi.org/10.60496/S6B88F

[11] Zooniverse. [Accessed 06-06-2024]. Zooniverse Citizen Science Projects. https://www.zooniverse.org/