WORKSHOP REPORT

3RD SMD AND ETD WORKSHOP ON AI AND DATA SCIENCE: LEAPING TOWARD OUR FUTURE GOALS

NASA GODDARD SPACE FLIGHT CENTER
MARCH 21-23, 2023

3rd SMD and ETD Workshop on A.I. and Data Science: Leaping Toward Our Future Goals

STEERING COMMITTEE:

- Mark Carroll – GSFC mark.carroll@nasa.gov
- Chris Neigh – GSFC christopher.s.neigh@nasa.gov
- Dan Crichton – JPL daniel.j.crichton@jpl.nasa.gov
- Manil Maskey – MSFC/HQ manil.maskey@nasa.gov
- Ed McLarney – LaRC edward.l.mclarney@nasa.gov
- Shubha Ranjan – ARC shubha.ranjan@nasa.gov
- Ryan McGranaghan – JPL ryan.m.mcgranaghan@jpl.nasa.gov
- Barbara Thompson – GSFC barbara.j.thompson@nasa.gov
- Michael Johnson – GSFC michael.a.johnson@nasa.gov
- Victoria Da Poian – GSFC victoria.dapoian@nasa.gov
- Marlo Maddox – GSFC marlo.maddox@nasa.gov

Report Lead: Mark Carroll, Data Science Group Lead, co-lead AI Center of Excellence NASA Goddard Space Flight Center

Writer and Editor: Peter Williams, PBW Consulting and Process Matters Group

Executive Summary

Throughout NASA, Artificial Intelligence (AI) is profoundly disrupting science, its methods, processes, and discoveries. On March 21-23, NASA's Science Mission Directorate (SMD) and Engineering and Technology Directorate (ETD) hosted the *Third Joint Workshop on AI and Data Science.* An overarching meeting goal was to share NASA's AI strategy for science and hear from the science community about embracing new technologies that leverage AI in science workflows and achieve new discoveries.

Workshop goals and agenda built on results from a 2022 *NASA-At-Work Challenge* that solicited ideas from across the agency about what is keeping NASA from achieving the next step in AI. It was those ideas that guided and anchored this meeting, connecting the two efforts. Nearly 160 NASA employees or contractors from seven different NASA Centers participated in the workshop either in-person or virtually, with presentations from both NASA and Industry experts intended to inspire and inform attendees about current activities and plans, as well as about opportunities not yet widely embraced.

A central theme that emerged from the meeting was: **NASA's AI Community and Partners Are Ready for A Sustained, Collaborative Effort to Advance AI-enabled Science.** Important work has been accomplished in science data discovery, science data analysis, and instrument autonomy. Much of this work, however, has been performed in pilot projects with little opportunity to continue the work beyond the short duration of the pilot projects. Given the rate at which science and technology are evolving, and that AI is disrupting them, there is a need for change because NASA AI-enabled science cannot afford to allow good ideas to be left behind.

There were six primary takeaways from the meeting:

- Important AI projects are underway at each of the seven NASA Centers represented at the workshop, many as pilot projects and a few as partnerships across Centers.
- Achieving the full potential for advances in AI demands greater collaboration between NASA centers, with other government agencies and with private industry and academia.
- Continuing NASA investment in High-End Computing infrastructure, including on-premises and commercial cloud, and commitment to Open Science will accelerate current and future work.
- An expansion of efforts in ethical, explainable, and trustworthy AI is needed to overcome the "fear" of AI that is hindering further uptake of AI in the broader community.
- AI-supported autonomy is fundamental for current and future missions to maximize science returns from missions in deep space and from those generating huge volumes of data.
- Ongoing support for training and learning in the rapidly changing ecosystem of AI is essential for our workforce to continue making advances and to help recruit and retain that workforce.

There is both depth and breadth to the AI and Machine Learning (ML) activities underway across all science divisions and engineering at NASA. The community is primed and ready to make great near- and long-term advances. These advances will be facilitated by retaining talented scientists and engineers and recruiting the next generation. The greatest strength of NASA is the science and engineering workforce's understanding of missions and data generated from those missions, because that understanding is essential to designing and implementing investigations of large volumes of data. The future of AI is upon us; it is time to take the next leap.

Table of Contents

Cover shows images of AI-related science and technology used by NASA or produced by NASA scientists.

From bottom to top:

Aurora seen from space: *NASA*
Crop and flood waters: NASA/Lauren Dauphin/USGS
Solar flare: *NASA/SDO*
Curiosity's dusty selfie: *NASA/JPL-Caltech/MSSS*
Cyclones of color at Jupiter's north pole: *NASA/JPL-Caltech/SwRI/MSSS*
Family portrait of the Jovian System: *NASA*
Illustration of neutron stars colliding: *Sonoma State Univ./A. Simonnet; NASA*

Introduction

Today, the science and technology of Artificial Intelligence (AI) is changing fast, affecting every NASA Directorate, every Center, every project. Throughout NASA, AI is proving to be a catalyst for scientific advancements both directly and indirectly. Directly, researchers analyzing scientific datasets are unlocking discoveries about the Universe not possible only a few years ago (Figure 1). Indirectly, AI-enabled improvements to High-End Computing and advancements to innumerable instruments are making science more efficient and powerful.

Tomorrow, the story of AI will be more complex and nuanced. It also will be more intimidating to some, just as it will be more inspirational to others. Today is a moment to prepare, to leap into that future.

This document reports on the Third Joint Science Mission Directorate (SMD) and Engineering and Technology Directorate (ETD) Workshop on Artificial Intelligence and Data Science held March 21-23, 2023, at Goddard Space Flight Center, Greenbelt, Maryland (agenda). The overarching goal for the meeting was to foster collaboration across centers and share NASA's AI strategy for science and hear feedback from the science community on how they can embrace new technologies to implement AI, including Machine Learning (ML), in their workflows to advance analyses for new scientific discoveries.

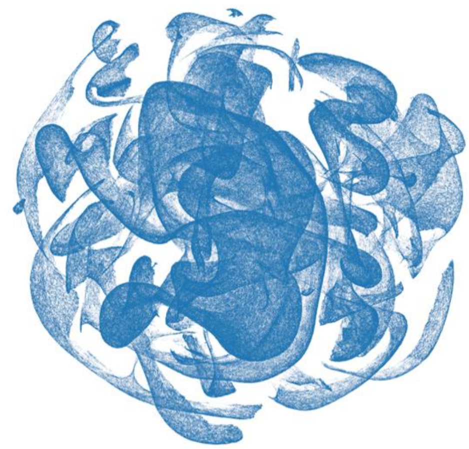

Figure 1. A two-dimensional projection of the high-dimensional space of TESS light curve representations. Image credit: Brian P. Powell, NASA Goddard.

That goal and the workshop agenda also incorporated results from a 2022 *NASA-At-Work Challenge* that solicited ideas from across NASA about what is keeping the agency from achieving the next step in AI, from leaping into the future. It was those ideas from NASA's workforce that guided and anchored this meeting.

The workshop was attended by nearly 160 NASA employees or contractors from seven different NASA Centers, with presentations from both NASA and Industry experts intended to both inspire and inform attendees about current activities, future plans, and opportunities for growth. A separate report contains summaries of each talk.[1] Five (5) concurrent focused discussions occurred as part of the meeting to allow participants to explore a particular topic in detail, identifying issues and opportunities and bringing both back to the larger group.

This report, however, responds to a central theme that emerged during the meeting:

NASA's AI Community and Partners Are Ready for A Sustained, Collaborative Effort to Advance AI-enabled Science.

[1] The document with summary notes of all workshop sessions is available upon request and will be posted on a pending NASA webpage together with this report.

The start of this effort is captured here, a leap towards future goals, a call to action. While important AI-related work has happened and is happening at NASA, there are opportunities for improvement and change. Participants at the workshop identified the opportunities and needs for change documented here. Still other opportunities and needs are seen in results from recent assessments of community members, and there are likely others not yet part of the discussion.

There is a difference, however, between recognizing that a need for change exists and having a shared understanding of changes to consider. A shared understanding of change can be seen as most immediately promising and we seek how to make those happen. These are questions about how members of our scientific community choose to work together and what is needed to work together well, these questions require commitment to continuous learning, as today's needs for change inevitably evolve into tomorrow's.

This report lays out a way forward, a way to build a strategy collaboratively to address these and similar questions, a way that is as responsive to NASA's AI community as it is to the moment. It sets the stage for a collaborative, systematic approach, an approach that is intentional and invites NASA's AI-community and its partners to participate, to join in this leap towards future goals.

Important NASA AI Work Has Happened and Is Happening Now

Throughout NASA, AI is profoundly disrupting science, its methods, processes, and discoveries. AI-informed scientific methods, like analyses to interpret data or algorithms to compress it, expand knowledge about crops and drought, about oceans and air, about planets, stars, and galaxies.

AI is disrupting the scientific process itself as *Large Language Models* (LLM) provide opportunities (and new challenges) for generating literature reviews and summaries of published work which may facilitate common terminologies within and across disciplines. The promise of LLM is faster, more efficient literature reviews allowing Principal Investigators to focus more on investigations and allowing team members to accomplish more in less time. In reality, we see mixed results with LLM and in many cases the summary outputs have been described as "hallucinations" that do not track with any known published research. The community expressed interest in these technologies together with a healthy skepticism of results until those are thoroughly vetted.

These changes, and many others, are the groundwork for profound new discoveries, like those made possible by AI-led advances in climate and ecosystem modeling, digital twins, and Foundation Models[2] (Bommasani et al., 2021). These are the projects at the forefront of science, the ones that will attract and retain the amazing employees of tomorrow.

---

[2] A *Foundation Model* for AI is understood to have three primary qualities: (1) it is pre-trained on a comprehensive dataset and capable of being used for multiple downstream tasks, (2) it substantially reduces the downstream effort of building AI applications, including the need for large, labeled training datasets, and (3) it is capable of capturing emergent behavior within the system being modeled, meaning it can learn to recognize behavior that is not the behavior targeted by the scientists.

Part of this current and transformative disruption in science stems from the increasingly integrated lifecycles of science/research, data/models, and infrastructure/technology. Science at NASA today must intentionally engage all three areas in a complex, integrated system of exploration and discovery.

In addition, a movement beyond interdisciplinary research is happening, what the National Academy of Sciences describes as *convergent thinking* and *convergence*, increasingly essential to addressing complex problems, expanding discoveries, and meeting edge challenges (Council, 2014). Such thinking demands an open, inclusive culture and a collaborative approach to problems.

In this way, convergence represents an expanded form of interdisciplinarity in which bodies of specialized knowledge comprise "macro" domains of research activity that together create a unified whole. When integrated effectively, these convergent macro domains offer the possibility of a new paradigm capable of generating ideas, discoveries, methodological and conceptual approaches, and tools that stimulate advances in basic research and lead to new inventions, innovations, treatment protocols, and forms and strategies of education and training (Council, 2014, p. 21).

NASA Divisions doing AI-related work have largely distinct, complementary approaches, although some are similar and some important partnerships have been happening. That work happens at multiple NASA Centers and engagement with industry and academia supports those efforts. Some AI methods are more widely used and familiar, others less so.

Clearly, important work has happened and is happening, yet NASA's approach to AI to date also raises important matters of science. For example, are some AI methods promising, yet poorly explored? If so, who takes on the exploration? Is the role of serendipity too great for the weight of the moment? Is the role of intentionality too little?

Opportunities For Improvement and Needs for Change

NASA's AI work has opportunities for improvement, including many presumably not yet recognized. Some opportunities relate to the capacity of NASA's AI community to work together, including knowing more about who is working on similar questions or using similar methods, to communicate and coordinate better, and to leverage investments across projects and programs.

There also are opportunities to do more by partnering with others, like private industry and academia, including existing opportunities not widely enough known. Similarly, there are opportunities to make NASA datasets, models, software, and algorithms more easily and more readily available, especially to those who have not benefited from previous access.

And there are opportunities for making invaluable AI-related training available more widely throughout the current AI community and to those who might be welcomed into it, building on objectives and work of NASA's Transform to OPen Science (TOPS) initiative. No one person can assimilate all the skills required to take full advantage of all aspects of AI, or any science. Hence, continuing and expanding access to training is essential to success and collaboration is the key to great advances.

One specific problem for NASA's AI community is the frequency with which good ideas fail to transition from successful proof-of-concept and prototyping to production and mission deployment. This is the graveyard between Technology Readiness Level (TRL) 4 and TRL 5 (Lavin et al., 2022). Too much of the work to date has struggled to cross this gap, as though it were more like a chasm. So far, reasons given for this struggle are more often anecdotal and less well substantiated, yet each of those reasons is an opportunity for improvement worth exploring. More to the point, this overarching problem presents a clear need for change because failing to cross that gap allows too many good ideas to perish. Yet, this is an opportunity to identify Best Practices that could become invaluable throughout the agency, across Directorates, Divisions, and projects.

Another problem is that many technical AI-products are often too difficult for domain scientists and researchers to interpret. This suggests a disconnect between data science products and domain science needs. What is least clear is who owns the problem, which suggests another needed change to clarify the problem, as well as to clarify who will have what role in addressing it.

Opportunities for addressing and improving this problem are already in motion, including the growing emphasis on *Explainable AI* to make AI more transparent and, thus, more trustworthy and likely to be adopted in critical science domains (Saeed & Omlin, 2023). The more AI is transparent, the more ethical it is and the less intimidating, all leading to better science and more valuable discoveries.

Perhaps most reassuring, however, is that there are opportunities to address these needs and others in ways consistent with the collaborative, convergent thinking that is so promising. Doing so would be invaluable because alignment between the thinking that inspires a community and the actual ways members of that community work together can prevent *cognitive dissonance*, can prevent a disconnect between values and work environment, a disconnect that can contribute to lower employee morale and retention, less constructive work environments, and less realization of scientific opportunities.

These are non-trivial problems that AI and the NASA AI-enabled science community can help address, either directly or indirectly. As an example, NASA's AI-enabled science community might find a way to establish an appropriate level of standardization, perhaps even including somewhat standardized ways of working together to maintain or adjust those actual standards, ways that avoid going too far and producing rigidity that would be anathema to NASA's scientific, innovative, and exploratory missions. Directly, AI can support standardization and benefit directly from it. Indirectly, being proud of having good, substantiated AI standards done well will lead to greater professionalization and retention of NASA employees who work on AI-science.

The most promising opportunity, then, would be to design and implement a systematic approach to build and support NASA's AI community of practitioners and domain scientists, partners and colleagues, an approach that nurtures discovery and encourages innovation within and across scientific domains, within and across NASA Centers, among and between NASA partners.

NASA’s AI Community is Ready for a Sustained Effort

Throughout the 2023 workshop, NASA’s AI community members expressed a readiness for a sustained effort to realize opportunities for improvement and address needs for change. Community members recognize these are not solely technical problems, nor are they merely structural problems. Rather, there is a common realization that seeing these challenges as one or the other will prevent the community from successfully navigating what Gartner has called the “AI-Hype Cycle” [3]. The challenges of AI-enabled science are increasingly understood as *socio-technical problems*, meaning ill-defined problems with social and technical dimensions (i.e., “wicked” problems[4]) (Baxter & Sommerville, 2010).

Advancing AI-enabled science will require social decisions about technical matters. NASA’s AI community members are ready because their science and NASA’s Core Values demand it (Figure 2). They are looking for standardization of workflows, of methodologies, of data, making each easier to find and use. They also are looking for shared repositories for data, models, workflows, training materials, and literature. They are looking for a shared literacy about AI, even if not every community member can become equally expert on all aspects. And they are looking for more common terminology across disciplines and domains, between Centers and with partners.

Figure 2. NASA's Core Values https://www.nasa.gov/careers/our-mission-and-values

(Photo Credit: NASA/Ben Smegelsky)

Yet, they also realize that *standardization* may not have a standard meaning and, if done poorly, it risks interfering with the very science that drives them. Done well, however, appropriate standardization promises to support more Open-Source Science that is also more powerful, just as it will support more inclusive and collaborative science that also drives discovery and innovation. For example, data accessibility is central to Open Science, yet centralized, decentralized, and hybrid approaches each offer different advantages.

This suggests the sustained effort for which NASA’s AI community is most ready would be one that is intentional and strategic. Yet, what that might look like could vary greatly, underscoring the ill-defined nature of the problem and, thus, that there is no single best answer.

Part of the solution likely could include the next iteration of NASA’s *Transform to Open Science* program (TOPS), an iteration that could help produce the sustained support the AI-enabled science community seeks. Building from lessons learned through the effort to transition to Open Science and Open Data, TOPS might evolve to provide continuous training opportunities, more

[3] Link: https://www.gartner.com/en/articles/what-s-new-in-artificial-intelligence-from-the-2022-gartner-hype-cycle

[4] *Wicked problems* are distinguished from *tame* ones by ten primary characteristics, including that there is no definitive formulation, no stopping-rule, no true or false solutions, only solutions that are good or bad, and often only to a varying degree. See: Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a General Theory of Planning. *Policy Sciences*, *4*, 155-169.

direct funding for AI opportunities and research, a platform for developing and disseminating more standardized resources, and greater access to collaborative resources including the commercial cloud.

Another part of the solution is to include data science as a design feature in future missions. This is being done on some levels but must be enhanced to take full advantage of the investments that NASA makes in its missions. Two examples include Autonomy and Data Management. AI applied to data management must include better ways to solve the problems of sending large amounts of data back to Earth from not only Earth observing missions but also missions to the solar system and beyond. Autonomy is when we take our AI systems to the next level to operate within a set of boundaries and rules. **Workshop participants noted that an effort is needed to identify contained use cases in preparation for use at a larger scale in future missions.**

The question, then, is how to find an appropriate approach, one that benefits from the best ideas available about what is working well and for whom, what gets in the way and why, what is most likely to set NASA's AI community and its partners up for success well into the future.

Finding the Right Approach into the Future

The problem or need of establishing a systematic, intentional, sustained effort by NASA's AI community is ill-defined, meaning there are multiple ways of understanding or framing the problem, of seeing or stating the need that should be addressed. Because of this, there are multiple ways of addressing the need or solving the problem. And this means it is a problem of the organization and community, a *socio-technical* problem, not purely a technical one.

Ill-defined social or organizational problems must be addressed in a collaborative way if the result is to be successful, collaborative, and consistent with convergence thinking, all as defined by the community members themselves. This is because, given that the need or problem is ill-defined, it is the community and its members whose insights must design the approach, must inform the thinking, must be willing to help realize and live with the outcomes, and must define success. Moreover, by defining the problem and need collaboratively, community members also will be more willing to live with that definition because they helped craft it and, thus, see it as appropriate.

A community that does not share an understanding of the need or problem to address is especially unlikely to find a shared way to address it. But, by reaching consensus about the problem and need, the community also is more likely to reach consensus about a solution and about what a successful solution would look like, including who would have what role in implementing it. Because the goal is to have a collaborative AI-enabled science community with the capacity to do more than it does today, community members must have opportunities to engage in turning that abstract, yet inspirational ideal into a concrete, intentional strategy that can serve as the basis for actual mission deployment.

Finding the right approach is of great importance because NASA's AI community faces its own gap between proof-of-concept and mission deployment, between TRL-4 and TRL-5. The now proven concept is that NASA's AI community would be well served by a *Knowledge Management*[5] based approach to a community of practice, one that brings people and programs together with information and knowledge using technology and tools (Figure 3).

The gap is between concept and a mission-deployed, and long-term need for a robust, sustainable AI effort. How the community chooses to cross that gap has the potential to positively influence every other effort, every other project, and every other program faced with the challenge of crossing similar ones.

Finding the right approach also matters because internal consistency is vital between the *process* by which the system is designed and the *ideals* for which it is designed. NASA's AI community seeks to attract the best ideas and the most commitment. That means the eventual system "outcomes" must be consistent with Open Science, with convergence thinking, and with collaborative ideas. It also means the process must be equally consistent.

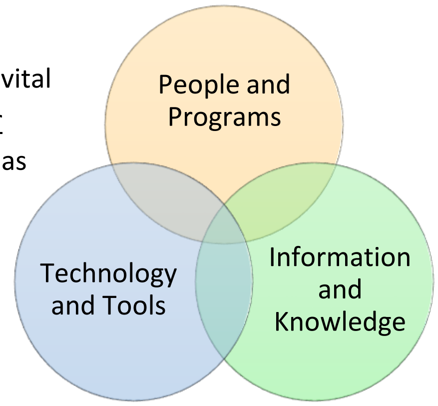

Figure 3. NASA's AI Community would be well served by a Knowledge Management approach to an intentional Community of Practice if it can cross the gap to deployment (Graphic: Peter Williams).

Perhaps most importantly, unless NASA's AI community members are committed to the process, their commitment to any outcome is less likely, meaning needed standards, metadata, communication forums, and more will lack the commitment—and support—essential for success. One measure of the right approach, therefore, is that it attracts and earns commitment to the process and, thus, builds commitment to the outcomes.

The way to do this is for NASA's AI community members to work together to clarify "needs" before moving to "solutions/tools." This is because a clear, shared understanding of a *need* makes it easier for the community to recognize and reach consensus on how to meet it. Clarifying "questions" and "goals" must happen before moving to "answers" and "actions". This is how to find the right approach. By inviting and encouraging NASA's AI community members to participate in a collaborative process design, they will set the stage for the process and products to be equally collaborative and convergent.

Leaping Towards Our Future Goals

NASA's AI community is prepared for a sustained effort towards more AI-driven scientific discoveries through data science. That is the community's overarching goal, the one powering it into the future. There is a pent-up energy ready for release, ready to spring, ready to leap towards this goal.

The community's goals are as inspirational as aspirational. They include a faster AI-driven scientific process, through literature reviews, hypothesis generation, project and program management, and data

[5] Knowledge Management is a sound framework for information management and science: https://www.tlu.ee/~sirvir/Information%20and%20Knowledge%20Management/Framework%20for%20IKM/about_this_learning_object.html

analyses. They include emboldening data management experts to solve increasingly important problems as the volume and complexity of datasets grow all but exponentially, as well as to leverage AI as an invaluable tool for science, focusing scientific exploration, separating noise from signal, and raising the probability of discoveries otherwise not even possible. And they include faster and better AI-driven data labeling, producing technically valuable results while exploring methodological questions of scientific interest.

The goals also include growing a stronger AI-supported scientific community by building community capacity to do more science and to keep up with the explosion of data. **AI has the potential of playing at least three key roles**:

1. As a *catalyst* for the community, AI-enabled science demands interdisciplinary teams and, thus, naturally attracts an interdisciplinary community of interest and practice, entirely consistent with convergence (Figure 4).
2. As a *lever*, AI techniques can help meet some of the AI-enabled science community needs, like building a common language or improving metadata.
3. And, as a *recruitment and retention tool*, NASA AI-enabled science will attract talent, inspire people, and retain professionals because NASA endeavors are simply compelling.

Some of the more inspirational goals for NASA's AI community might include possible workflow improvements, like those that might facilitate crossing the gap between TRL-4 and TRL-5, as well as questions about what standards would be appropriate and what a workflow process for standardization might look like.

Equally inspirational goals could be about what trustworthy AI might mean and how AI might help manage AI and maintain trustworthiness, perhaps by looking for anomalies before any breach of trust occurs. AI also might help manage High-End Computers, perhaps integrating Machine Learning (ML) with Development and Operations (i.e., MLOps), making code more efficient and effective, saving time, money, and energy, eventually integrating quantum computing once that technology becomes mature and deployable, certainly an inspirational goal as well.

Figure 4. Nano-catalyst (Credit: European Space Agency)

By joining in the effort to leap towards these goals, community members invest in themselves as much as in the community. For example, clearly not everyone can do or know everything, so how the community is built determines what the community can do, and what the community can do will allow community members to concentrate on their strengths and interests, doing more together while helping themselves. **Similarly, a stronger community will mean NASA Centers become less siloed, making the work, products, and knowledge less opaque, less obtuse, and more easily and readily shared.**

Seeing Our Possible Future

The future for NASA's AI-enabled science community is inspirational and compelling. While details remain to be crafted, rough contours are coming into focus. The 2023 workshop participants have a vision of that future, even if important details are fuzzy. They see a sustained, systematic, intentional approach to AI that builds on existing cross-Center partnerships to allow community members to do more, especially in five specific areas:

- Working within and across Centers, programs, and projects (collaboration and partnerships),
- Sharing models and algorithms (tools),
- Sharing trained, common, or commonly structured datasets (labeled, training, validation),
- Knowing who is doing what (community-building and efficiency),
- Leveraging expanded partnerships with private industry and academia.

They see a community coordinating resources and creating efficiencies, developing *Best Practices* and *Standard Operating Procedures*, and constantly improving through training and support, AI-literacy, mentoring, and coaching. They see *TOPS* evolving, building on what has worked well and serving as an invaluable vehicle to develop, test, and deploy new ideas, a vehicle to pursue continuous improvement.

NASA's AI-enabled science community also sees a future in which daunting topics receive the right attention. Ethics, for example, is a great concern in AI-enabled science with at least two distinct areas:

(1) Ethical AI, which is about characteristics of AI systems and technology, and
(2) AI Ethics, which are practices guiding and anchoring professional AI-related work.

By wrestling with the matters of Ethical AI and AI Ethics, NASA's AI-enabled science community sees a future where trust in AI is well-placed, where speculative concerns—like enfeeblement, deception, emergent goals, or power-seeking behavior—are addressed empirically to reduce, bound, and manage them (Hendrycks & Mazeika, 2022).

To address the importance of ethics, a vision for NASA's AI-enabled science future might include best practices in both these areas, practices drawn from multiple fields, consistent with a convergent approach. **An *AI Code of Conduct* might inspire community members to identify best practices and document some of the more important practices as conduct or principles to emulate**. Perhaps most importantly, the community would consider what it might mean to approach ethics as a process or culture, not a destination or product.

And they see a future in which NASA's AI-enabled science community identifies and shares best practices for recruitment, hiring, and retention, making better and broader use of the federal Position Description that exists for data scientists (Figure 5).

This possible future all but certainly would see greater integration and professionalization of *Data Science* and *Data Scientists*, recognizing that these two needs could be somewhat in tension: Integrating is to embed, like an equal, indistinct member of a team, while professionalization often implies elevating above today's status, still a member of a team, yet meaningfully distinct as a profession. Nevertheless, workshop discussions made it clear the community sees both as needed, suggesting that the possible future would be based on a *search* for a convergent both/and approach, rather than a debate about the right either/or choice.

In this possible future, AI-ready datasets and Foundation Models are more widely understood and shared in ways that leverage appropriate consistency while protecting appropriate flexibility. The promise of Foundation Models hinges on AI-ready datasets, but whether a dataset is AI-ready may depend on the domain, scale, or AI method. In the future, this tension is embraced and addressed, perhaps with an approach that either parallels or is integrated with that of Technology Readiness Levels (TRLs) and addresses exceptionally practical questions, like who is expected to do the work of producing AI-ready datasets, those who produce the data or those who need the data? Who is responsible? Who determines when a dataset is AI-Ready? AI-Ready for whom?

Figure 5. LEGO "Build The Future" Activity (Photo credit: NASA/Bill Ingalls)

Lastly, this possible future might reflect a NASA AI-enabled science community that has wrestled with prioritization, looking honestly at Edge Technologies because some hold future possibilities, while others offer immediate, near certain benefits. Quantum Computing and some other edge technologies, for example, are not advanced enough to support current needs and need a few years before they are useful. Similarly, access to the Cloud is desired by the community, yet deciphering how best to go about it in a cost-effective manner is still an opaque challenge. And AI-specific hardware, like chips, boards, and even full High-End Computers/Supercomputers, are available, yet not integrated with workflows as broadly as appears needed.

A Summary of Opportunities for the Future:

- NASA Collaboration: AI projects are underway at each of the seven NASA Centers represented at the workshop, and science can be accelerated by encouraging and facilitating collaboration across Centers.
- AI Ethics and Ethical AI: Ethical, explainable, and trustworthy AI must be encouraged and, where possible, codified into an AI code of conduct to ensure everyone knows what is expected and AI-transparency becomes possible.
- Autonomous Science: Autonomy, enabled by AI, is essential to maximize science returned by current and future missions in deep space and those generating huge volumes of data.
- Infrastructure: Investments in infrastructure, including on-premises and commercial cloud, are essential to future discoveries because collocated data and compute are vital for advanced AI algorithms and Open Science.
- Training and Learning: Opportunities for training and learning in the rapidly expanding ecosystem of AI are essential to enable our workforce to continue making advances.
- Recruitment and Retention: Expanding recruitment efforts for trained data scientists to be partnered with existing pool of scientists is another crucial investment need.
- Partnerships: NASA can’t do it alone, hence partnerships between NASA and external partners, including private industry, academia, other federal agencies, and international partners, are critical for making large advances in AI and AI-enabled science.

References

Baxter, G., & Sommerville, I. (2010). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, *23*(1), 4-17. https://doi.org/10.1016/j.intcom.2010.07.003

Bommasani, R., Drew, Adeli, E., Russ, Sydney, Michael, Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Jared, Demszky, D., Donahue, C., . . . Pang. (2021). On the Opportunities and Risks of Foundation Models. *arXiv pre-print server*, 214. https://arxiv.org/abs/2108.07258

Council, N. R. (2014). *Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond*. The National Academies Press. https://doi.org/doi:10.17226/18722

Hendrycks, D., & Mazeika, M. (2022). X-Risk Analysis for AI Research. *arXiv pre-print server*. https://doi.org/10.48550/arXiv.2206.05862

Lavin, A., Gilligan-Lee, C. M., Visnjic, A., Ganju, S., Newman, D., Ganguly, S., Lange, D., Baydin, A. G., Sharma, A., Gibson, A., Zheng, S., Xing, E. P., Mattmann, C., Parr, J., & Gal, Y. (2022). Technology readiness levels for machine learning systems. *Nature Communications*, *13*(1), 6039-6058. https://doi.org/10.1038/s41467-022-33128-9

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a General Theory of Planning. *Policy Sciences*, *4*, 155-169.

Saeed, W., & Omlin, C. (2023). Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. *Knowledge-Based Systems*, *263*, 110273. https://doi.org/https://doi.org/10.1016/j.knosys.2023.110273

Appendix: Attendance Summary — 3rd NASA SMD+ETD AI Summit

ATTENDANCE SUMMARY	All three days (totals of individuals)	
	Subtotal	Total
Virtually Only (#)		66
In-Person Only (#)	71	
In-Person/Virtually (#)+	22	
In-Person (total)*	93	93
Total Attendees**		159
Check count		159
NASA Employee or Contractor***		134
JPL Employee		12
Industry		6
Other		7
Total Attendees		159

+ "In-Person/Virtually" = People who attended some in the room and some on WebEx.

* "In-Person (total)" = sum of "In-Person only" and "In-Person/Virtually".

** "Total Attendees" = sum of "Virtually Only" and "In-Person (total)".

*** Seven NASA Centers, including JPL, were represented: (1) Ames Research Center, (2) Goddard Space Flight Center, (3) Jet Propulsion Laboratory, (4) Langley Research Center, (5) George C. Marshall Space Flight Center, (6) Lyndon B. Johnson Space Center, and (7) John H. Glenn Research Center.