



Reports of the Association for the Advancement of Artificial Intelligence's 2024 Spring Symposium Series

May 23, 2024 | Reports, Symposium Reports, Updates

By Jessica Coates, Mononito Goswami, Takashi Kido, William Lawless, Xinyu (Rachel) Li, Christopher J. MacLellan, Andreas Martin, Siddharth Srivastava, Reinhard Stolle, Keiki Takadama, Pulkit Verma, Jie Yang, Melo-Jean Yap

The Association for the Advancement of Artificial Intelligence's 2024 Spring Symposium Series was held at Stanford University in Stanford, California, March 25-27, 2024. There were eight symposia in the spring program: Bi-directionality in Human-AI Collaborative Systems, Clinical Foundation Models Symposium, Empowering Machine Learning and Large Language Models with Domain and Commonsense Knowledge (AAAI-MAKE 2024), Federated Learning on the Edge, Impact of GenAI on Social and Individual Well-being, Increasing Diversity in AI Education and Research, Symposium on Human-Like Learning, User-Aligned Assessment of Adaptive AI Systems. This report contains summaries of the workshops, which were submitted by some, but not all, of the workshop chairs.

Bi-directionality in Human-AI Collaborative Systems (S1)

The symposium on "Bi-directionality" combined two tracks: "Human-Centered Computing in the Age of AI" and "Risk Perceptions and Determinations In Collaborative Human and AI-Based Systems." Two introductions were given, one for each track. The first introduction was given by Jie Yang, Andrea Tocchetti, Lorenzo Corti, and Marco Brambilla, representing TU Delft and Politecnico di Milano; the second introduction was given by Ranjeev Mittu, Information Technology Division, Naval Research Laboratory, in Washington, DC. A total of nineteen speakers representing universities, businesses, and government agencies from around the world participated in our symposium. The

symposium involved invited talks, presentations of accepted papers, panel discussions, and speed talks. The biographies and abstracts for all of the speakers in the program are to be found on our supplementary web page:

<https://sites.google.com/view/bidirectionality2024/program>.

The substance of the symposium addressed the challenges in creating synergistic human and AI-based autonomous systems-of-systems. Recent advances in generative AI techniques (e.g., LLMs) have exacerbated the growing concerns associated with AI, held by researchers and the public alike, such as the risk, trust, ethics, and safety to the users and to the public from the operations of autonomous machines/AI alone in open situations. These concerns present major hurdles in the development of verified and validated engineered systems involving bi-directional pathways across the human-machine barrier; in this context, bi-directionality means understanding the design and operational consequences that the human may have on machine agents and the effects that machine or AI agents may have on humans. Current discussions on human-AI/machine interactions are unresolved or fragmented, focusing either on the impact that AI or machines may have on human stakeholders (including the relevant human factor considerations) or potential ways of involving humans or machines in computational or physical interventions (e.g., data annotations, human-machine behavior interpretations, operator-machine interventions). We believe the challenges associated with human-AI/machine collaborative systems cannot be adequately addressed if the underlying challenges associated with bi-directionality are not fully identified and taken into consideration.

Speakers at our symposium addressed a wide range of topics. The highlights of the talks included what human-machine bi-directionality meant; the explainability of decisions by machines and humans with a shared language, and the greater risks that may arise when no shared language exists; the real and perceived risks, trust, ethics and safety of human and machine teams operating autonomously in the field; how shared human-machine awareness and mental models may reduce risks; systems designs and engineering; and the assurance, testing and evaluation of autonomous human-machine systems. In particular, several talks addressed the current challenges in Natural Language Processing (NLP) research, highlighting the need for Human-Centered methodologies to make meaningful progress in NLP.

The symposium has been followed up with a call for revised and extended chapters to be published later this year by Elsevier, co-edited by Lawless, Mittu, Sofge, and Brambilla.

The co-chairs for the “Human-Centered” track were Jie Yang, Andrea Tocchetti, Lorenzo Corti, and Marco Brambilla; the co-chairs for the “Risk Perceptions” track were William

Lawless, Ranjeev Mittu, Don Sofge, and Hesham Fouad. This report was written jointly by Jie Yang and William Lawless.

Clinical Foundation Models Symposium (S2)

Foundation models are quickly emerging as powerful tools to solve a variety of biomedical challenges, such as clinical text generation and summarization, radiograph analysis, disease prediction, etc. These models are characterized by their ability to solve multiple prediction tasks across diverse domains and have great potential to streamline healthcare administration, reduce costs, enhance accessibility, and ultimately improve the quality of patient care. However, many questions remain unanswered: (1) What qualifies as a clinical foundation model? (2) Which healthcare challenges can be effectively addressed by clinical foundation models, and which remain beyond its scope? (3) What are the challenges associated with training and applying these models in a healthcare context? (4) How can such models be best integrated with the routines of healthcare professionals? This symposium convened researchers and practitioners from artificial intelligence and healthcare fields across academia, industry, and government to deliberate on these key challenges, exchange insights, and identify promising pathways for future research and practical implementations.

The symposium featured seven keynote talks, three panel discussions, four technical demos and tutorials; ten contributed talks, and two poster sessions. Keynote speakers from academia, healthcare institutions, and industry delivered talks covering diverse topics pertinent to clinical foundation models. Prof. Frederic Sala from the University of Wisconsin-Madison started off the symposium with insights on data-efficient foundation models, from their pre-training to adaptation. Mr. Vivek Natarajan from Google Research presented their work on scaling healthcare through medical LLMs and AI systems like Med-PaLM and AMIE. Dr. Gilles Clermont highlighted the potential of foundation models, particularly in modeling high-frequency time-series data, for clinical decision support in critical care. Professor Jimeng Sun shared insights into developing Generative AI technologies to aid clinical trial design and operation. Dr. Hoifung Poon from Microsoft Research Health Futures discussed recent progress in generative AI for precision health, covering biomedical LLMs, multimodal learning, and causal discovery. Mr. Mononito Goswami introduced MOMENT, the first open-source suite of large, pre-trained foundation models for time series, while Dr. Erhan Bas from GE Healthcare presented their work in grounding large vision language models. Panel discussions delved into challenges, opportunities, and future directions in developing clinical foundation models and integrating them seamlessly into clinical practice. Additionally, technical demos and tutorials by start-ups such as Nixla, Snorkel, and Gradient AI, along with Amazon AWS AI

Lab, showcased recent work and products ranging from time-series foundation models to generative AI solutions for drugs and adverse event extraction and innovative platforms integrating proprietary medical foundation models with user-facing data extraction and generation capabilities.

At the symposium, 32 papers were featured, with ten selected for contributed talks. While the majority of papers centered around topics related to foundation models, such as their pre-training, fine-tuning, and evaluation in clinical settings, as well as considerations regarding fairness, transparency, data curation, and benchmarking, there were also contributions focusing on traditional machine learning methods for clinical tasks and position papers offering a comprehensive perspective on the application of foundation models in healthcare. The papers related to foundation models spanned various data modalities encountered in clinical scenarios, including text, images, time series, and genomics data. During the two poster sessions held on consecutive days, authors presented their work to a large audience, sparking meaningful discussions. All accepted papers are available for review at https://openreview.net/group?id=AAAI.org/2024/Spring_Symposium_Series/Clinical_FMs.

In summary, the AAAI 2024 Spring Symposium on Clinical Foundation Models was highly successful in bringing more than 81 participants together to discuss challenges and opportunities for foundation models, and machine learning for healthcare, making it one of the largest symposiums within the AAAI 2024 Spring Symposium Series. Attendees expressed overwhelmingly positive feedback, prompting plans for its continuation as a workshop in upcoming conferences.

Artur Dubraski chaired the event, with Mononito Goswami and Xinyu (Rachel) Li serving as co-chairs. Other key organizers included Gilles Clermont, Su-In Lee, Cecilia Morales, Tristan Naumann, Frederic Sala, and Jimeng Sun. This report was co-authored by Mononito Goswami and Xinyu (Rachel) Li.

Empowering Machine Learning and Large Language Models with Domain and Commonsense Knowledge (AAAI-MAKE 2024) (S3)

The AAAI Spring Symposium on Empowering Machine Learning and Large Language Models with Domain and Commonsense Knowledge convened a rich mosaic of minds. This assembly of researchers, practitioners, and industry professionals explored the interplay between machine learning, knowledge engineering, and the rapidly evolving

field of large language models (LLMs). Central to the discussions were the challenges of enhancing AI systems with trustworthiness, interpretability, and commonsense reasoning, a trio of topics deemed critical for the future of AI.

Ron Brachman, former director of the Jacobs Technion-Cornell Institute and Professor of Computer Science at Cornell University, kicked off the symposium with a keynote on “AI with Common Sense—The Bigger Picture.” Brachman underscored the necessity of imbuing AI with human-like common sense and argued for a holistic approach to cognitive AI systems. In his analysis, common sense is located somewhere between the “Fast” and “Slow” systems from Kahneman, exhibiting the speed and simplicity of the “Fast” system as well as being cognitively penetrable like the “Slow” system. Brachman proposed several ideas for the implementation of common sense and traced them back to Minsky (frames), Schank (scripts), and McCarthy (Advice Taker).

Later on, Noah Goodman, Professor of Psychology and Computer Science at Stanford University and a Research Scientist at Google DeepMind, presented “Reasoning in humans and machines.” Goodman’s talk offered an intriguing comparison between machine intelligence and human cognitive processes. He showed how cognitive models of reasoning correspond to some techniques in LLMs, such as chain-of-thought reasoning. In a “self-taught reasoning” approach, systems can learn rationales in a reinforcement learning feedback loop.

The symposium also paid tribute to the late Douglas B. Lenat, a pioneer in artificial intelligence, through the reflections of Andreas Martin, Professor of Applied Artificial Intelligence at the FHNW University of Applied Sciences and Arts Northwestern Switzerland, and Edward A. Feigenbaum, Professor Emeritus of Computer Science at Stanford University. Their joint keynote, “In memoriam of Douglas B. Lenat (1950–2023),” celebrated Lenat’s contributions to AI.

Alison Gopnik, Professor of the Graduate School in the Department of Psychology at the University of California, Berkeley, delved into “Large Language Models as Cultural Technologies: Truth versus Transmission.” Gopnik’s research on children’s learning and understanding of the world provided a unique lens through which to view LLMs. As an approach to building commonsense reasoners, she argued for a combination of causal learning, which maximizes information gain, with reinforcement learning, which focuses on receiving a specific reward.

On the symposium’s second day, Henry Lieberman from the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL) challenged attendees with “The Next Grand Challenge in AI — Making Better Mistakes.” Lieberman introduced the distinction between “charming” mistakes and “bizarro” mistakes and argued that understanding this

distinction is a good starting point for exploring methods to incorporate common sense into intelligent systems.

The concluding keynote by James McClelland, Lucie Stern Professor in the Social Sciences at Stanford University, was titled “Incorporating explanation-based reasoning into models of human and machine intelligence.” Based on an introspective analysis of what constitutes a thought, McClelland offered an intriguing vision for future AI development.

Throughout the symposium, sessions highlighted advancements in trustworthy AI, commonsense reasoning evaluation in LLMs, human-centered AI, and the merging of knowledge engineering with machine learning, among other pivotal topics. These discussions underscored the vast potential and pressing need for AI systems that can navigate the complexities of human knowledge and reasoning, moving toward a future where AI can better reason and interact with the world in a human-like manner.

The diverse range of topics and the depth of discussions at AAI-MAKE 2024 illustrated the vibrant, ongoing dialogue at the intersection of machine learning, knowledge engineering, and LLMs. As the field continues to evolve, the insights and collaborations fostered at this symposium will undoubtedly shape the trajectory of AI research and development for years to come.

Andreas Martin (chair), Pedro A. Colon-Hernandez, Maaïke de Boer, AURORA Gerber, Knut Hinkelmann, Pascal Hitzler, Jane Yung-jen Hsu, Yen-Ling Kuo, Xiang Lorraine Li, Thomas Schmid, Paulo Shakarian, Reinhard Stolle, and Frank van Harmelen served as the organizers of this symposium. This report was written by Andreas Martin and Reinhard Stolle and edited by Pedro A. Colon-Hernandez and Yen-Ling Kuo.

Federated Learning on the Edge (S4)

Traditional Artificial Intelligence (AI) models predominantly rely on centralized computing architectures, limiting their potential in scenarios where real-time decision-making on low-latency devices is required. AI on The Edge has emerged to overcome these limitations, allowing AI algorithms and models to be deployed directly on edge devices, such as sensors, IoT devices, and autonomous systems. This shift in computation distribution reduces latency, improves responsiveness, and aims to enhance privacy, security, and bandwidth consumption. The next iteration for Edge AI is to allow devices to learn together and collaborate under a unified system architecture. The Federated Learning (FL) computational paradigm can facilitate this transition. This symposium invites academia, industry, and government researchers to explore Federated Learning on The Edge and its unique challenges and opportunities. The symposium will invite submissions of extended abstracts (to be developed into four-

page manuscripts). In addition, the symposium will host invited keynote and session speakers. Overall, this symposium will offer a unique opportunity for participants from various backgrounds and agencies to engage in lively discussions, network with peers, and foster collaborations to advance and guide research and development for Federated Learning on The Edge. No formal report was filed by the organizers for this symposium.

Impact of GenAI on Social and Individual Well-being (S5)

The AAAI 2024 spring symposium on “Impact of GenAI for Social and Individual Well-being” explored the profound intersections between generative AI and human well-being, focusing on both individual and societal impacts, ethical considerations, and practical applications.

The AAAI 2024 spring symposium on “Impact of GenAI for Social and Individual Well-being” was held at Stanford University, California, from March 25th to 27th. The emergence of generative AI (GenAI) has led to a profound intersection between society and human well-being. Although GenAI’s potential enhancements to our daily lives are immense, they also present unique challenges. As we further incorporate GenAI into societal frameworks, the emphasis should not be solely on technological prowess or economic benefit. It is equally crucial to ensure ethical considerations, such as fairness, transparency, accountability, and the protection of privacy and security. For example, in healthcare, the generative models used in diagnostics must be both precise and interpretable. The data that these models operate on must be comprehensive, representing various cultures, ages, genders, and geographical areas to accurately mirror the complexities of our diverse societies. The impact and potential of GenAI in creative arts, education, and journalism are expected to be equally profound and challenging. Given GenAI’s significant influence, it’s essential to establish ethical boundaries in this era. This symposium explores this multifaceted topic from two primary perspectives.

The first perspective, “Individual Impact of GenAI on Well-being,” aims to elucidate the mechanisms and issues to consider when designing AI and GenAI for personal well-being. In this context, the focus excludes societal aspects. Topics may include efficiency in individual work enhancement, personalized medical care, support in learning and education, new forms of entertainment, and privacy concerns. The discussion should center on how AI and GenAI can enhance opportunities for individual well-being, with an emphasis on understanding the emotional and quality-of-life implications of these technologies.

The second perspective, “Social Impact of GenAI on Well-being,” intends to highlight the mechanisms and issues to consider when incorporating societal aspects into GenAI for well-being. Topics may involve changes in employment structures due to automated AI, preventing the worsening of social inequalities, the potential to enhance the quality of health and medical treatment, the risk of misinformation spread, and ethical debates on AI’s judgment criteria and values. Exploring the social impact of GenAI on well-being is anticipated to shed light on both the potential benefits and risks of AI and GenAI. We must also explore ways to prevent machines from adopting human bias, ensuring fairness, and producing socially responsible outcomes.

We welcome both technical and philosophical discussions on the individual and social “Impacts of GenAI” on well-being, particularly in the realms of ethical design, machine learning software, robotics, and social media (though not exclusively). Topics such as interpretable forecasts, responsible social media, beneficial robotics, combating loneliness with AR/VR, and promoting personal health were pivotal in our discussions. This symposium aimed to share the latest advancements, current challenges, and potential applications related to social responsibility for well-being. Evaluations of digital experiences and insights into human health and well-being should be encouraged.

Our symposium included 27 technical presentations over two-and-a-half days. Presentation topics included (1) the challenges of GenAI for Well-being; (2) Generative AI: Harmful and Ethical aspects; (3) Generative AI: Well-being and Learning; (4) Generative AI: Analysis; (5) Fair, Diversity, Equity, and Inclusion; (6) GenAI’s application for well-being; (7) AI Health Agents; and (8) well-being for patients and healthy persons.

For example, Takashi Kido, the Advanced Comprehensive Research Organization of Teikyo University in Japan, presents challenges to Social and Individual Well-being. Oliver Bendel, School of Business FHGW in Switzerland, presented how GenAI can foster well-being in self-regulated learning. Melanie Swan, University College London in the United Kingdom, presented AI Health Agents, illustrating pathway2vec, ReflectE, category theory, and longevity. Jin Yamanaka, Fujitsu Research of America in the United States, presented an evaluation of large language models with RAG capability from a perspective of robot behavior planning and execution. Faye-Marie Vassel, Stanford University in the United States, presented the psychosocial impacts of generative AI harm. Andy Skumanich, Innov8ai Inc. in the United States, presented models of tracking mal-info in social media with AI/ML tools to help mitigate harmful GenAI for improved societal well-being. Michelle Nie, Sciences Po in France, presented the biggest threat to democracy today by Artificial Intelligence.

Our symposium offers participants unique opportunities to develop new ideas through innovative and constructive discussions among researchers from diverse backgrounds.

This highlights the significant interdisciplinary challenges that will guide future advances in the AI community.

Takashi Kido and Keiki Takadama co-chaired this symposium and wrote this report.

Increasing Diversity in AI Education and Research (S6)

The symposium titled “Increasing Diversity in AI Education and Research” stood as a pivotal milestone in the ongoing pursuit of inclusivity within the artificial intelligence education landscape. Led by Mary Lou Maher, PhD (UNC Charlotte) and Jessica Coates (Spelman College), this event brought together a diverse cohort of scientists, educators, and industry leaders.

Over the course of three dynamic days, the symposium hosted an impressive lineup of speakers, including notable figures like Kamau Bobb (The Georgia Institute of Technology), Karen Colbert (Keweenaw Bay Ojibwa Community College), Charity Freeman (CSTA Board Chair), and Ebony McGhee (The Johns Hopkins University), and Anshul Sonak (Intel). These visionary leaders set the tone for the symposium, challenging attendees to confront issues of diversity, equity, and inclusion in AI.

In addition to thought-provoking talks, the symposium featured ten paper presentations that featured topics on (i) AI at Minority-Serving Institutions (MSIs), (ii) AI for good, and (iii) Addressing Digital Inequality. All authors hailed from MSIs or historically excluded communities. These presentations underscored the importance of amplifying underrepresented voices in AI research and education.

The symposium’s impact extended well beyond its duration, laying the groundwork for future initiatives. Insights from organizations such as the National Science Foundation (NSF), The Integration of Learning & Synthesis (TILOS), and Inclusion through Virtual Intervention and Technology Enrichment (INVITE) enriched discussions on navigating the funding process and fostering diversity in AI.

Acknowledgments to all participants and the rest of the Organizing Committee (in alphabetical order): Nate Derbinsky (Northeastern), Bonnie Dorr (University of Florida), Judy Goldsmith (University of Kentucky), Naja Mack (Morgan State University), Jamie Payton (Temple University and Invite AI), Jodi Reeves (National University and TILOS), Mehran Sahami (Stanford), Neelu Sinha (Fairleigh Dickinson University), Melo-Jean Yap (The Johns Hopkins University), and Clement G. Yedjou (Florida A&M).

Co-chairs Mary Lou Maher, PhD (UNC Charlotte) and Jessica Coates, PhD (Spelman College) led this symposium. This report was authored by Jessica Coates and Melo-Jean Yap.

Symposium on Human-Like Learning (S7)

The AAAI 2024 Spring Symposium on Human-Like Learning aimed to bring together researchers working across diverse approaches and methodologies to (1) identify key learning capabilities that humans exhibit and that machine-learning systems currently do not, (2) explore ongoing and proposed research into how to create systems with these human-like learning capabilities, (3) determine how to evaluate these kinds of systems, and (4) discuss their potential benefits. In pursuit of these aims, participants at the symposium presented and discussed 21 talks and ten posters.

The symposium kicked off with a talk by Pat Langley (from the Institute for the Study of Learning and Expertise) on his “Gauntlet of Human-Like Learning,” which emphasized several key characteristics of human learning that might be targeted for machine learning research. One of the capabilities that was repeatedly touched on throughout the symposium was data efficiency and the ability to learn from small rather than big data. For example, talks by Maya Malaviya (from Stevens Institute of Technology) and Ilia Sucholutsky (from Princeton University) explored the development of systems that—like humans—can learn from extremely limited data; they even presented approaches for learning classifiers in situations where there is less than one datum per class.

Beyond specific capabilities, the talks and posters also explored the role that research on representation, decision-making, analogy, concept and task learning, and language modeling might play in better understanding human learning and in developing human-like learning systems. For example, Kenneth Forbus from Northwestern University and Irina Rabkina from Barnstorm Research Corporation introduced the idea that analogical systems might serve as a fundamental tool—a kind of Swiss army knife—for creating human-like learning systems. Daniel Weitekamp, from Carnegie Mellon University, explored how multiple specialized learning mechanisms can produce more efficient skill learning than a single mechanism. Lastly, Peter Lindes from the Center for Integrated Cognition presented ideas on how systems might be architected to support more human-like language and concept acquisition. These are just a few illustrations of topics discussed at the event. A full list of presentations and abstracts can be found at <https://humanlikelearning.com/aaai24-ss/>.

A central question discussed throughout the symposium was, what does it mean for a system to support human-like learning? One key insight was the importance of a focus on learning, not just performance. This led to the conclusion that human-like learning systems cannot just be evaluated based on final performance alone—evaluations must measure aspects of the learning process, too, such as learning rate and data efficiency. A

second crucial insight was that “human-like” refers to understanding and replicating functional human capabilities rather than mimicking how humans instantiate these capabilities (i.e., by being brain-like). Thus, a human-like learning system might simultaneously exhibit human learning capabilities—such as being able to learn incrementally and continually—while also being able to process data at speeds and scales that far exceed what a human can do.

During the discussion on broader impacts, participants highlighted several potential benefits of human-like learning techniques, such as more efficient learning with smaller energy and carbon footprints, improved data efficiency and the ability to learn from small data, the ability to efficiently learn continually and on the fly—letting systems stay up to date with the latest information—and supporting improved human-AI interaction by behaving and learning in human relatable and interpretable ways. Across all our discussions, we determined that there are clear opportunities for more future research into human-like learning.

Christopher J. MacLellan, Douglas Fisher, Ute Schmid, and Randolph M. Jones were members of the organizing committee. This report was written by Christopher J. MacLellan.

User-Aligned Assessment of Adaptive AI Systems (S8)

This symposium aimed to address research gaps in assessing the compliance of adaptive AI systems (systems capable of planning/learning) in the presence of post-deployment changes in requirements, user-specific objectives, deployment environments, and the AI systems themselves.

The symposium featured invited talks focusing on formal verification by Prof. Sanjit A. Seshia from the University of California, Berkeley, and Prof. Sriram Sankaranarayanan from the University of Colorado Boulder, on using AI for assessment by Prof. Sriraam Natarajan from the University of Texas Dallas, and Prof. Kamalika Chaudhuri from the University of California, San Diego, and Meta AI.

The symposium brought together representatives from academia, industry, and the government to discuss the critical aspects of evaluating and regulating adaptive AI systems that can learn and update themselves after deployment. It tackled unresolved issues in assessing the compliance of these systems when faced with evolving requirements, user objectives, operating environments, and changes to the AI systems themselves through updates and continual learning.

Several contributed talks also focused on issues like aligning AI systems with the

designer’s intent, the user’s intent, or the user’s expectations. A few talks concentrated on various ways to collect data about requirements and formalize them using technologies like formal logic and large language models. The symposium provided a platform to explore novel techniques for conceptualizing, expressing, and enforcing regulations for these highly adaptive AI systems that can evolve in unanticipated ways post-deployment.

A few contributed talks even explored approaches for combining neural networks with various “symbolic” approaches like hierarchical goal networks and dual-system theory of mind models. This raises questions about whether AI systems should have built-in rigid interpretations coded by designers or if they should develop more flexible, interactive interpretations through experience.

There was debate around using large language models (LLMs) to provide explanations of other LLM-based AI systems, with some concerns that this could lead to degenerate or uninformative descriptions if not done carefully. But the common consensus was that researchers need to be careful in trusting output of these LLMs.

Many of the examples and applications discussed stemmed from transportation domains like self-driving vehicles, delivery drones, medical applications, and routing optimization problems. However, there was also considerable focus on medical and public health use cases, such as AI systems for personalized patient treatment regimes.

Throughout these talks, a common thread was addressing the inherent dissimilarities between an AI system’s initial training/design objectives set by developers and the practical, real-world objectives and preferences of the end-users, which may evolve over time. Resolving this interpretation tension is crucial for ensuring user trust and regulatory compliance as these adaptive AI systems learn and update themselves during deployment.

Pulkit Verma, Rohan Chitnis, Georgios Fainekos, Hazem Torfah, and Siddharth Srivastava served as co-organizers of this symposium. This report was written by Pulkit Verma and Siddharth Srivastava.

Authors

Jessica Coates is a Lecturer in the Department of Biology at Spelman College

Mononito Goswami is a Ph.D. student at Carnegie Mellon University’s Robotics Institute.

Takashi Kido is a professor at Teikyo University in Japan.

William Lawless is a Professor of Mathematics and Psychology, Paine College, Augusta, GA

Xinyu (Rachel) Li is a Ph.D. student at Carnegie Mellon University’s Robotics Institute.

Christopher J. MacLellan is an Assistant Professor in the School of Interactive Computing at Georgia Institute of Technology.

Andreas Martin is a Professor of Applied Artificial Intelligence at the FHNW University of Applied Sciences and Arts Northwestern Switzerland.

Siddharth Srivastava is an associate professor at the School of Computing and Augmented Intelligence, Arizona State University.

Reinhard Stolle is the Director of the Business Unit Mobility at Fraunhofer Institute for Cognitive Systems, Munich, Germany.

Keiki Takadama is a professor at the University of Electro-Communications in Japan.

Pulkit Verma recently graduated with a PhD from Arizona State University.

Jie Yang is an Assistant Professor in the Web Information Systems group at TU Delft, The Netherlands.

Melo-Jean Yap is a Senior Education Research Consultant in the Center for Teaching Excellence & Innovation at The Johns Hopkins University.