iSAT Snapshot

NSF AI INSTITUTE FOR STUDENT-AI TEAMING

Fall 2025

In Brief . . .

A quick look at our activities this quarter!

1. PI D'Mello and Strand 1 Co-Lead, Jacob Whitehill (WPI), participated in panel discussions at the "Inaugural International Workshop on Rethinking - Children's Automatic Speech Recognition For Education" in Buffalo, NY. D'Mello participated on the topic of "Children's ASR in Action at National AI Institutes and Centers" and Whitehill on the topic of "Data Drives True AI, but Where do We Stand Today?"

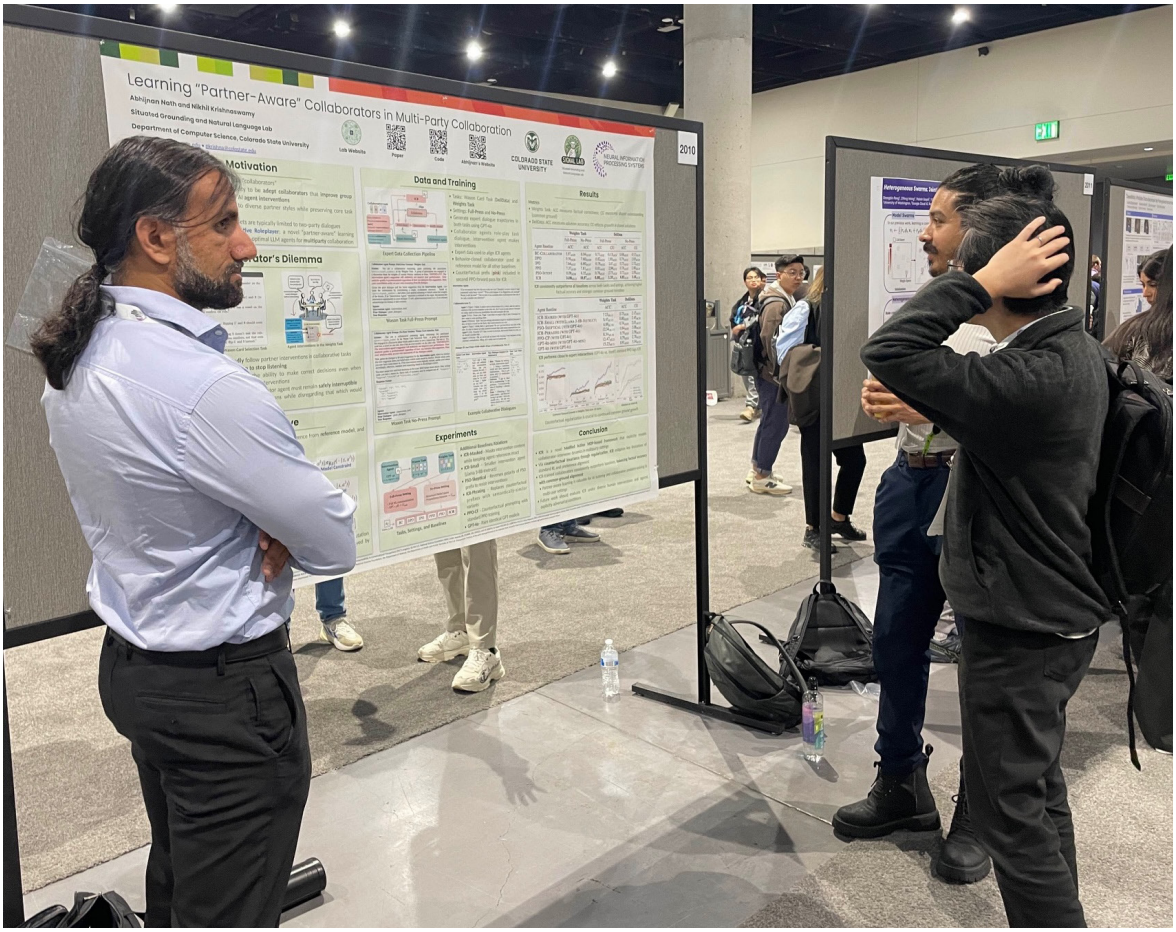
2. D'Mello and Executive Director, Peter Foltz, attended the Summit for AI Institutes Leadership (SAIL) meeting in Reston, VA. D'Mello helped deliver the opening remarks of the conference along with members of two other AI Institutes. He also served on a plenary panel on Interdisciplinary Research Collaborations: Success Stories at the conference.

3. NSF iSAT postdoc, Ekta Sood, and PhD student, Kelechi Ezema, also attended the SAIL conference, which welcomed early career scholars this year. They both presented their work at the Graduate Student & Postdoc Posters Showcase at the conference.

4. D'Mello and Foltz also participated in a pre-conference workshop of the five NSF AI for Education Institutes. They presented an update on research at NSF iSAT and participated in discussions identifying key research threads where cross-institute collaboration would have the greatest impact.

5. Foltz served on the advisory board for the National Academy of Education development of a roadmap for generative AI in K-12 Learning and Assessments.

6. Foltz co-led a half-day workshop at the NCME AI in Measurement Conference on using generative AI for assessing and supporting learning. Strand 3 members Greg Benedis-Grab, Jeffrey B. Bush Tayne, and Strand 3 co-lead William R. Penuel attended the National Science Teachers Association conference in Minneapolis, MN. They led seven sessions on our storylined curriculum units, community agreement routines, and responsible innovation for educators.



Strand 1 member, Nikhil Krishnaswamy, presenting at the 39th Annual Conference on Neural Information Processing Systems, San Diego (CA).

Feature: Shaping What's Next

Highlights from Our Design Retreat in Boulder, Colorado

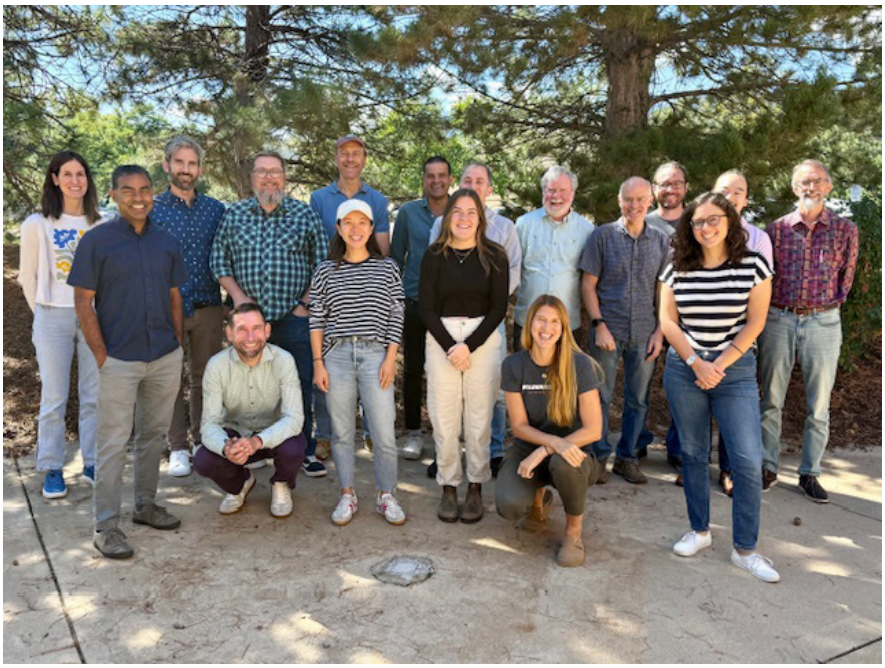
This September, iSAT's senior personnel came together at the University of Colorado in Boulder for a 1.5-day design retreat that marked the start of Phase II of our research project. The retreat was both an opportunity to reflect on what we learned in Phase I and a chance to develop new ambitious goals to shape iSAT's next five years. With members from all strands represented, the retreat brought about lively exchange, creative thinking, and a deep commitment to our collective vision of transforming classrooms into knowledge building communities through responsible innovation with AI.

Shared Goals, Values, and Vision

We began Day 1 with roundtable discussions in cross-strand groups over breakfast to update everyone on how each strand is planning to contribute to our larger goals; this set the stage for deeper conversations to follow. We then focused on discussing our values and commitments around our within- and cross-disciplinary research efforts. The purpose was to identify what our collaborations should look like in practice and to align those aspirations with our broader goals. Facilitators guided attendees in coalescing around shared values and norms, agreeing on internal processes such as communication and decision-making, generating ideas for mentorship and connection across strands, and identifying themes that could guide institute culture moving forward. One of the most powerful metaphors to emerge was that of a classroom "on fire"—not in a literal sense, but in the sense of a space that sparks joy, ignites passion, and sustains curiosity among students. This evocative phrase captured the collective aspiration to create classrooms that not only engage but also inspire.

Building Convergence Around the AI Partners

The retreat then shifted to one of its central activities: building convergence around a new range of AI Partners. In preparation for the retreat, each strand created imaginative narrative documents depicting possible small group interactions with AI in future classrooms orchestrated by teachers. During the session, attendees first gathered with their strands to review the scenarios from the other strands, tease out points of convergence, and generate proposals that could shape Phase II AI Partners. The goals of the activity were multifaceted: to identify key features of AI designs and classroom practices that might support expansive collaboration in the next five years while holding a longer-term horizon in mind; to make explicit the commitments guiding design work through responsible innovation so that values could be shared and transparent; to articulate the classroom, school, and professional supports necessary to bring each scenario to life; and to begin sketching initial design descriptions that could later be refined in conversation with students and teachers. Over lunch, facilitators took on the task of synthesizing these broad ideas into a smaller set of high-level groupings of design features. This synthesis was critical, as it created a bridge between the imaginative scenarios of the morning and the design sessions of the afternoon. During the afternoon, we also dedicated a session to reviewing video footage from previous classroom implementations of our existing AI Partners to give our team a more immediate window into how student groups and teachers interacted with our AI Partners in real-time. The clips highlighted a wide range of collaborative behaviors, responses to AI support, and adaptations that teachers made to curriculum routines.



Members of iSAT's Senior Personnel at the Design Retreat in Boulder, Colorado.

Seeing these moments together helped surface new design questions and inspired several directions for future tool development. The team then engaged in extensive design sessions where each institute member could select one of the emerging high-level design features (multimodality, trustworthiness, etc.) and work in small groups to develop concrete design concepts for the AI Partners and supporting structures (e.g., teacher professional learning). This was the most generative moment of the day: groups brainstormed metaphors and sketches for novel AI tools and classroom supports. Each group then presented its concepts back to the larger team, which concluded the day.

Initial AI Partner Designs & Future Planning

Day 2 opened with informal discussions over breakfast before attendees re-convened for another design session. This time, the focus was on the strands meeting within their own groups to assess how they could contribute to the high-level designs identified during the previous day. This helped each strand to locate itself within the broader vision of iSAT in Phase II while also considering how to collaborate with others to achieve cross-strand convergence and goals. The retreat concluded with a whole-group session on design convergence and future planning. Here, the group turned explicitly toward the upcoming Phase II Strategic Implementation Plan (SIP). We identified four major clusters of work to serve as initial launching points towards developing the SIP in the next couple of months.

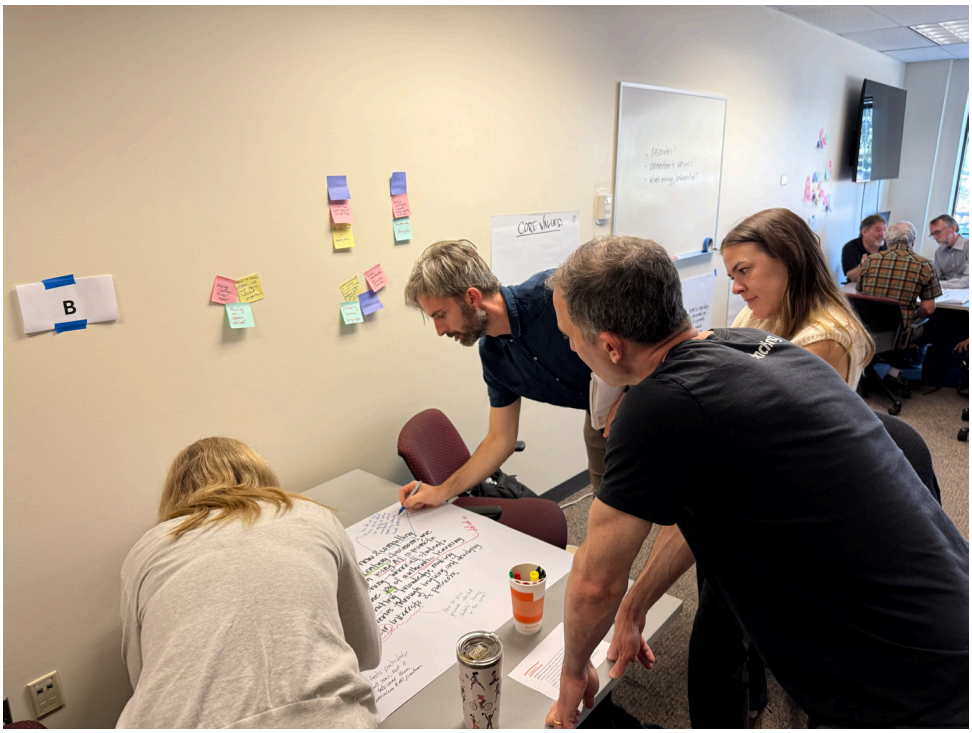
AI Partners: enhancing the Community Builder (CoBi) AI Partner and generalizing the Jigsaw Interactive Agent (JIA) from Phase I, developing the Connection Finder (which surfaced from design sessions with youth in Phase I), and designing novel AI tools for curriculum implementation support.

Cross-Cutting Themes: addressing essential issues around AI usage such as privacy and trust, transparency in how AI systems function, multimodality, configurability of AI partners, generalizability across contexts, and building pathways for students and educators to develop and build AI literacy.

Methods: creating supports for rapid prototyping and experimentation in real-world classrooms (including Wizard-of-Oz techniques) alongside literature reviews to situate designs in broader scholarly conversations.

Nexus: scaling curriculum units and the AI Partners, extending AI partners beyond curriculum-specific boundaries, advancing the science of student-AI teaming, and deepening partnerships with research communities, schools, districts, and other communities of practice.

Beyond these formal outcomes, what stood out most about the retreat was the atmosphere and the sense that it was a family reunion of sorts. Attendees described the experience as energizing, noting how rare and valuable it is to spend extended time thinking together across disciplinary boundaries. The metaphor of a classroom "on fire" applied as much to our vision for students and the retreat itself: it was a reaffirmation of NSF iSAT's mission and a demonstration of how interdisciplinary collaboration can yield both creativity and convergence. With Phase II now underway, the retreat provided both the compass and the north star to carry this important work forward.



Interdisciplinary collaboration at the September design retreat in Boulder, CO.

Strand 1: Building for Real-Time Multimodal Collaboration

Fall Updates

Strand 1 research aims to address the foundational question: What advances in Artificial Intelligence (AI) are needed to facilitate collaborative learning conversations in classrooms? This semester's research focused on Processing Multimodal, MultiTalker Speech, Multimodal Nonverbal Behavior, and Interpreting Content.

Multimodal, MultiTalker Speech Processing

During the fall semester, the team has constructed a prototype real-time feedback mechanism with which users can correct the outputs of an automatic speaker diarization system (e.g., ECAPA-TDNN) as well as automatic speech recognizer (e.g., Whisper, Google cloud speech). The system consists of a multi-stage large language model (LLM) based workflow that recognizes a "wakeword" spoken by the user ("Hey CoBi, it wasn't the moisture Tensor, it was sensor"), identifies the desired correction (e.g., re-assigning an utterance attributed to speaker X to another speaker Y), and then modifies the speech and diarization transcript accordingly. Moreover, they have developed a simple user interface to visualize the running transcript and corrections in real-time. Finally, they have conducted experiments, using the Augmented Multi-party Interaction (AMI) Meeting dataset, showing that the proposed system substantially reduces diarization error rates (DER) by 9.92% and speaker confusion error by 44.23%. In addition, the team has designed a novel architecture for multi-talker automatic speech recognition (ASR) to capture simultaneous speech.

Multimodal Nonverbal Behavior Processing

Strand 1 team members have also completed a perception system for detecting non-verbal and verbal behaviors and extracting events during small group collaboration. They developed and refined evaluation metrics for the two-dimensional (2D) perception pipeline (including gaze tracking, pose detection, and action), designed to assess both the performance of the perception system itself and its effectiveness in supporting downstream tasks. The team continued developing the multiparty, multimodal perception pipeline with a focus on defining evaluation metrics to more systematically assess the performance and interpretability of the perception systems. Most recently, Strand 1 team members evaluated gaze performance using our Nonverbal Interactions in Collaborative Learning Environments schema (NICE) and are currently assessing posture and pointing performance.

Interpreting Content

Finally, the team is in the ideation phase of a project to use natural language processing to analyze students' dialogue to support a new AI partner in the works: the Connection Finder. So far, the team has watched and discussed iSAT's Moderation Unit videos to examine idea exchanges between students; to think about what information would be useful to surface to teachers (in discussion with Strand 3); and to think about how ideas progress with respect to the curriculum learning objectives.

Strand 2: Investigating Interaction with and Trust in AI

Fall Updates

Strand 2 research aims to address the foundational question: What advances in theories, interaction paradigms, and frameworks are needed to orchestrate effective student and teacher interactions with AI Partners? This semester's research focused on Assessing Nonverbal Collaboration, Evaluating Mixed-Initiative Supports, and Understanding Trust in AI.

Assessing Nonverbal Collaboration

Over the past semester, Strand 2 team members have been finalizing research on nonverbal indicators of collaboration. A paper under review analyzes youth interactions within iSAT's Moderation Unit, examining how students' verbal and nonverbal behaviors functioned across two jigsaw tasks (where students become experts on one part of a topic and then teach their group members) of differing cognitive complexity. Considering both modalities provided a more complete picture of group dynamics, revealing clear differences by task. In these lessons, certain verbal behaviors and signals of nonverbal engagement aligned with on-task focus, while others were more common during off-task moments. These findings suggest that learning supports that scaffold shared attention and help coordinate tool use may help students engage more productively with both the task and with one another.

Evaluating Mixed-Initiative Supports

Strand 2 members have also been investigating how to evaluate mixed-initiative interactive AI supports. They are developing a human-centered evaluation framework for conversational AI. Using data where middle-school groups in the iSAT lab engaged in an open-ended Science, Technology, Engineering, Mathematics (STEM) jigsaw activity, the team compared real-time human subject matter expert (SME) interventions with post-hoc responses from an LLM-based version of the Jigsaw Interactive Agent (JIA).

Early results show that automated or pre-trained metrics (e.g., BLEU, BERTScore) fail to capture pedagogically meaningful dimensions of support, such as alignment with students' understanding or opportunities for 'contingent' scaffolding that is responding in-the-moment to students' understandings and needs. Human annotations reveal that the LLM often over-scaffolds with more structure provided even when the students demonstrated higher comprehension. This underscores the limitations of current evaluation practices and the need for learning-science-grounded methods to assess pedagogical appropriateness.

Understanding Trust in AI

Strand 2 has also led cross-strand iSAT discussions to identify systematic techniques for evaluating youths' reliance on, uptake of, and trust in mixed-initiative AI supports during classroom group work. This includes reviewing recent Human-Computer Interaction (HCI) literature and quantifiable metrics for assessing the "success" of AI supports in mixed-initiative contexts. The team is integrating these insights with work evaluating JIA to assemble a set of candidate measures for assessing classroom AI Partners. These multi-dimensional metrics will guide iterative design and evaluation and will be shared broadly to support systematic evaluation across the AI and education (AIED) research community.

Strand 2 has also begun two projects utilizing Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines designed to improve the reporting of systematic reviews. One project involves a PRISMA literature review of trust in individuals, technology, and AI in classroom settings to develop a preliminary Student AI-Teaming Trust Framework. Building on foundational work in interpersonal trust, trust in automation, trust calibration, and trust in AI, this effort recognizes that existing models do not directly translate to collaborative learning environments. Trust-building in classrooms—already complex without AI—creates an opportunity for new research on trusted AI for learning. This review and synthesis will serve as the foundation for this framework.

The second project (also using PRISMA review guidelines) examines how teachers and students can collectively assess trust in AI tools used in classrooms. This interdisciplinary review will examine how learning scientists have addressed the ethical implications of data collection and the counter-approaches developed in secondary classrooms (i.e., designs and pedagogies that researchers have taken to address these ethical issues). The resulting insights will inform co-design of an AI audit with school teachers and students for iSAT partners.

Strand 3: Creating Systems for AI-Enhanced Instruction

Fall Updates

Strand 3 research aims to address the foundational question: In what ways can co-design processes empower stakeholders to envision, co-create, critique, and apply AI learning technologies for their schools and communities? This semester's research focused on Analyzing AI Literacy Frameworks, Co-designing a New Generative AI and Art Unit, Redesigning Assessments, and Implementing AI-Enhanced Materials.

Analyzing AI Literacy Frameworks

Over the past semester, as part of a discovery and needs assessment study, Strand 3 conducted an initial scan and analysis of frameworks for AI literacy to help guide the revision of existing units and development of new units. In the scan, the team identified a multi-institutional effort from the Organization for Economic Co-operation and Development (OECD) as one that is likely to have significant policy influence, since it includes key researchers and groups like AI4K12 (whose 2020 framework iSAT has been using). The OECD (2025) draft framework identifies five domains of knowledge and four domains of practice: engaging with AI, creating with AI, managing AI, and designing AI. This fall, an interdisciplinary group of teachers, AI researchers, and learning scientists engaged in a systematic "unpacking" and analysis of the knowledge domains and specific statements about what students should know. In unpacking, the team identified key knowledge components that need to be addressed in instructional materials, as well as potential investigations and experiences that could help students develop that knowledge. In addition, Strand 3 has identified initial candidate knowledge goals that will be addressed across current (to be revised) and future units, so that all the knowledge goals in the materials will be addressed.

Co-Designing Generative AI and Art Unit

Strand 3 has also made significant progress this fall in engaging teachers to develop a new unit on Generative AI and Art using the new framework. The team started by leading a group of seven researchers and educators in the first part of a process of co-designing a storyline unit: analysis of learning goals, brainstorming of candidate design challenges to anchor the unit, and eliciting student perspectives on candidate design challenges. Team members also developed an AI ethics framework that will be integrated into units. Subsequently, teacher co-designers piloted an initial lesson within multiple classrooms, and developed a draft storyline for the unit. A key finding from this project has been the need to interleave learning activities with design activities, to support both learning sciences researchers and teachers in developing deeper understanding of AI models, especially those that are newer and may not be appropriate to teach to middle schoolers (e.g., diffusion algorithm to generate art).

Redesigning Assessments

As part of a review of existing iSAT units, Strand 3 also identified the need for the development of a more comprehensive assessment system for iSAT units that aligns with the new AI literacy framework, and that can be used to evaluate iSAT's AI-integrated curricular activity systems. To that end, Strand 3 is using an evidence-centered design approach to develop and test these assessments. So far, the team has identified initial assessment targets for units using the OECD framework and developed a set of design patterns for use in revising assessments. Design patterns include: (1) pre- and post-unit knowledge assessments, (2) short end-of-class formative assessments ("exit tickets"), (3) and performance tasks that will assess student fluency with designing, managing, and evaluating AI. In addition, the team has applied the OECD framework to redesigning pre- and post-unit knowledge assessments for the Sensor Immersion unit and administered them in pilot classrooms.

Implementing AI-Enhanced Materials

Finally, Strand 3 researchers have been analyzing the implementation of AI-enhanced instructional materials, focusing on both the spring implementations of the AI Partners in classrooms. Implementation research is a type of research that focuses on documenting the how and why of adaptation of materials in specific contexts; it is essential to addressing the "triple innovation challenge" presented by needing to teach about a novel subject (AI), use new tools (AI partners), and use new pedagogical approaches (storyline materials). One set of analyses Strand 3 has undertaken focused on identifying how teachers experienced the first of these challenges: learning to teach about a novel subject. Findings show that for emerging domains like AI, content learning cannot simply be a preface to implementation; it must be treated as a central, ongoing part of the design process itself, integrated deeply into the teacher experience. This could potentially be mitigated through integrating more "teachers as learners" opportunities during the co-design processes. More scaffolding is necessary to support teachers to fully implement the more advanced AI concepts beyond sensing, collecting data, and basic programming such as image recognition, machine learning, and neural networks. Teachers need more time to master these AI concepts in order to be more comfortable teaching them, especially given the fast-paced nature of AI in today's culture. Co-design offers one way forward to center the needs of all the stakeholders involved and to balance the needs that do not so easily align. Integrating more content focused "teacher as learner" moments into the co-design process can provide the necessary scaffolding to support teachers to feel comfortable enough to implement learning around complex AI concepts.

Nexus Hub: Broadening iSAT's Reach

Fall Updates

The iSAT Nexus is the institute's "Connector Hub", bridging the gap between AI research and the needs of society. This semester's work focused on Expanding K-12 Partnerships and Socializing the Science of Student-AI teaming.

Expanding K-12 Partnerships

Nexus presented seven sessions at the National Science Teaching Association (NSTA) Conference, establishing iSAT as a national leader in AI curriculum and Responsible Innovation routines. The team also advanced a scaling strategy by forming the Nexus Scaling Working Group and held its initial foundational meeting with key district partners to co-design the Nexus Scaling Plan. During a discovery phase, the team identified "Curriculum-Tethering" (the requirement for full unit adoption for the AI partners to be used) as the primary bottleneck to national scaling and is working on ways to "untether" the AI Partners so they can be used more broadly.

The Nexus Hub also expanded partnerships by collaborating with the Colorado Education Initiative (CEI) to expand iSAT's district partners. This partnership allows iSAT to scale its work beyond the initial cohort to rural and under-resourced districts across the state.

Socializing the Science of Student-AI Teaming

Nexus also made progress to socialize the science of student-AI teaming. This semester, the team focused on building internal capacity for outreach by meeting with iSAT strand leads to identify key workshops and research/industry partnerships to serve as Nexus strategic foci. Nexus community-building efforts included meetings with AI Curricula development organizations to identify potential collaborations on AI curricula and partners. iSAT members also presented on iSAT research at the National Science Foundation (NSF) Summit for AI Leadership workshop on AI4Ed, identified collaborations with other NSF education institutes, and co-led a workshop at the National Council for Measurement in Education (NCME) Conference on AI and Measurement on large language models for assessment and support of complex skills.

Reference:
OECD. (2025). Empowering learners for the age of AI: An AI literacy framework for primary and secondary education (Review draft). OECD.

New Members

Learn more about our members!

Meet Theodora Chaspari, Strand 2 member and Associate Professor in Computer Science and the Institute of Cognitive Science (ICS) at the University of Colorado Boulder.

Q: What does your research focus on?
A: My research focuses on modeling human cognition and behavior via multimodal signals, such as tone of voice and language use. Over time, this work has evolved from traditional signal-processing approaches to advanced multimodal machine learning. More recently, I have focused on ensuring that these AI systems are trustworthy via designing algorithms that safeguard personal information and avoid treating individuals differently.

Q: What is the coolest thing about your research?
A: One of the coolest aspects of my research is how interdisciplinary it is. I have had the opportunity to collaborate with experts across a wide range of fields; from clinical psychologists studying couples' therapy, to industrial/organizational psychologists exploring job interviews and space mission teamwork, to urban and landscape architects examining how people interact with their built environments. These collaborations allow us to tackle complex human-centered problems from multiple perspectives, which makes the work very exciting.

Q: What has been a turning point or defining moment in your career?
A: A defining moment came when I took my first signal processing class as a junior. It was the first time I saw how mathematics and pattern-recognition algorithms could come together to create technologies with real-world impact. That realization opened my eyes to the broader potential of computer science being not just as work that powers machines, but as a field that enables meaningful applications like speech recognition, assistive technologies, and even self-driving cars. It was the moment I knew I wanted to pursue research that connects computational methods with human-centered challenges.

Q: What do you like to do outside of work?
A: Outside of work, I love exploring new places and enjoying Colorado's beautiful landscapes!

Aaron Gluck is a PhD student at the University of Colorado Boulder and member of the Boulder Language and Social Technologies (BLAST) lab. Aaron is a Computer Scientist and Computational Linguist working in the areas of Natural Language Processing (NLP) and Machine Learning. He is currently developing a Knowledge-Seeking framework, which leverages research in the fields of Education and NLP/ML to encourage language models to actively participate in self-directed inquiry. Aaron's broad interests lie in the learning of both machines and humans, and what we can learn about one from the other.

Amanda Hernandez Sandate is a PhD student in the Computer and Cognitive Sciences program. Her research interests focus broadly on how computational models can deepen our understanding of human cognition and improve how artificial systems collaborate and communicate with people in complex environments. She received a B.Sc. in Cognitive Neuroscience and Computer Science from Brown University, where she contributed to research at the intersection of neuroscience and AI.

Carine Graff is a PhD student in computer science (natural language processing) at CSU. Her current research interests include automatic speech recognition and low resource languages. She holds an MS in computer science (artificial intelligence) from SMU and a PhD in translation studies from Kent State University. Carine has a background in language education and translation, with extensive experience teaching French and translation as an instructor and professor in the US and abroad.

Zachary Kaufman is a PhD student at the University of Colorado Boulder studying computer science, cognitive science, and neuroscience. His research is in human–robot interaction, human computer interaction, and the design of AI systems to support mental health and cognitive well-being. Zach has extensive experience bridging engineering and the social sciences, with a particular interest in applying robotics and AI to address mental health challenges and cognitive disorders.

iSAT Publications

Published / Accepted

Breideband, T., Bush Tayne, J.B., Reitman, J.G., Rose, S., Weatherley, J., Penuel, W.R., D'Mello, S.K. (2025). A Feasibility and Implementation Integrity Study of the Community Builder (CoBi): An AI-based Collaboration Support System in K-12 Classrooms. International Journal of Artificial Intelligence in Education.

French, D., Moulder, R., Ezema, K., Wense, K. v. d., & D'Mello, S. K. (2025). Linguistic Alignment Predicts Learning in Small Group Tutoring Sessions. Findings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP 2025), Suzhou, China.

He, X., and Whitehill, J. (2026). Survey of End-to-End Multi-Speaker Automatic Speech Recognition for Monaural Audio. Computer Speech and Language Processing.

Ko, M.-L. M., Penuel, W.R., Hoang, N., Cline, B., Heskett, C., Reiman, J.G., Watts, E., Chandler,C., Foltz, P., Dey, I., Eden, J., Dede, C. (2026). Learning from Implementation Research: How Curriculum, Professional Learning, AI Tools and Classroom Studies Can Support AI-Literacy. Symposium session accepted for presentation at the meeting of the American Education Research Association. April 8-12, Los Angeles, CA.

Nath, A., and Krishnaswamy. N. (2025). Learning "Partner-Aware" Collaborators in Multi-party Collaborations. Adv. Neural Info. Process. Systems.

Penuel, W. R., Philip, T. M., Chang, M., Sumner, T., & D'Mello, S. K. (2025). Responsible Innovation in Designing AI for Education: Shifting from Personalization to Collaborative Problem-Solving. Educational Researcher.

Submitted / Under Revision

Archival, Peer-Reviewed/Under revision Papers

Breideband, T., Bush Tayne, J.B., Benedis-Grab, G., Reitman, J..G., Rose, S., Watts, E., Clines, B., Penuel, W.R, & D'Mello, S.K. (Submitted). When no plan survives first contact: Enactment complexity in embedding multimodal, multiparty, multiuser AI into real-world classrooms. In Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems.

Breideband, T. and D'Mello, S.K. Intelligent Tools to Enhance Collaboration in Small Groups. [Manuscript submitted for book chapter]. In S.W.J. Kozlowski et al.. Computational Group and Team Dynamics: Forging an Interdisciplinary Science. Oxford University Press (OUP).

Chandler, C., Ezema, K., French, D., D'Mello, S.K. (Submitted). Validating NLP Models for Real-World Educational Impact: Ecological and Criterion Evidence from Tutoring Discourse. Association for Compuational Linguistics (ACL).

Foltz, P. W. Key Considerations for Assessing and Supporting Collaboration with AI. [Manuscript submitted for book chapter]. In M. Shermis & M. Xiong. The Role of AI in Assessment: Revolutionizing Education. Taylor and Francis.

Foltz, P. W., Sawaya, S. & Chandler, C. Assessing and supporting collaboration skills through Generative AI. [Manuscript submitted for book chapter] In M. C. Mayrath, D. Robinson, & J. Behrens (Eds.). The Handbook of Generative AI In Education. Springer.

Guan, Y., Trinh, V.A., Voleti, V., and Whitehill, J. (Submitted). MLLM-based Speech Recognition: When and How is Multimodality Beneficial? IEEE Transactions on Multimedia.

Ko, M.-L. M., Chandler, C., Rose, S., Clines, B., Watts, E., Reitman, J., & Foltz, P. W. (Submitted). Supporting Collaboration in Multi-Party Small Group Activity: A Design Research Study of a Jigsaw Interactive Partner. Learning Analytics and Knowledge (LAK) 2026 Conference.

Sawaya, S., Bush Tayne, J.B., Ezema, K., Breideband, T., Jacobs, J., & D'Mello, S.K. (Submitted). AI Tools that Celebrate rather than Evaluate Student Discourse during Learning. In Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems.

Zhu, Y., Bradford, M., Lai, K., Obiso, T., Venkatesha, V., Pustejovsky, J., and Krishnaswamy, N. (Submitted). Distributed Partial Information Puzzles: Examining Common Ground Construction Under Epistemic Asymmetry. Language Resources and Evaluation Conference (2026).

Non-Archival Talks/Posters/Presentations

Biddy, Q., Benedis-Grab, G., Hoang, N., & Bush Tayne, J.B. (Submitted). Humans in the loop: Teacher expertise and the limits of co-designing AI curriculum. In Proceedings of the 20th International Conference of the Learning Sciences.

Behrens, J. & Foltz, P. W. (2025). What you need to know to unlock your generative potential. Workshop presented at the National Council of Measurement in Education AI and Measurement Conference, October.

Bush Tayne, J.B., Benedis-Grab, G., Biddy, Q., Clevenger, C., Chang, M., Breideband, T., & Penuel, W.R. (Submitted). From design to dialogue: How classroom routines mediate reflective collaboration with AI partners. In Proceedings of the 20th International Conference of the Learning Sciences.

Bush Tayne, J.B., Santo, R., Hoang, N., Heskett, C., Ko, M.-L. M., Eden, J., Penuel, W.R., Biddy, Q., Campanella, M., Reitman, J.G., Breideband, T., D'Mello, S.K., Benedis-Grab, G., Sawaya, S., Jacobs, J., Dey, I., Malamut, J., Demszky, D., Swanson, H., Solola, I., Munir, R., Sinha, R., Nguyen, H., Simon, S. A., Shiwalia, B. M., & Makransky, G. (Submitted). Evaluating and assessing learning in AI-embedded environments: Perspectives across learning scientists. In Proceedings of the 20th International Conference of the Learning Sciences.

Bush Tayne, J.B., Hoang, N., & Dey, I. (Submitted). Investigating support for collaboration on the AI frontier: Mixed-methods study of middle-grades computer science classrooms. In Proceedings of the 20th International Conference of Computer Supported Collaborative Learning.

Foltz, P. W. & Sawaya, S. (2025). Toward more principled approaches for AI measurement of complex skills: Five Considerations. Talk presented at the National Measurement in Education AI in Measurement Conference.

Guan, Y., and Whitehill, J. (Submitted). Transformer-encoder trees for efficient multilingual machine translation and speech translation. International Conference on Acoustics, Speech, and Signal Processing, 2026.

He, X., Guan, Y., Paurana, B., Dai, Z., and Whitehill, J. (Submitted). Interactive Real-Time Speaker Diarization Correction with Human Feedback. International Conference on Acoustics, Speech, and Signal Processing, 2026.

Ko, M.-L. M., Worsley, M., Chang, M. A., Chandler, C., Rose, S., Clines, B., Watts, E., Weatherley, J., Foltz, P., Raju, R., Muddaluru, R., Bush Tayne, J., Benedis-Grab, G., Biddy, Q., Breideband, T., Penuel, W. R., Sumner, T., Heskett, C., Hoang, N., Eden, J., & D'Mello, S.K. (Submitted). Designing AI partner-embedded learning environments to support AI literacies. In Proceedings of the 20th International Conference of Computer Supported Collaborative Learning.

Ko, M.-L. M., Penuel, W.R., Hoang, N., Cline, B., Heskett, C., Reiman, J. Watts, E., Chandler, C., Foltz, P., Dey, I., Eden, J., Dede, C. (2026). Learning from Implementation Research: How Curriculum, Professional Learning, AI Tools and Classroom Studies Can Support AI-Literacy. Symposium session accepted for presentation at the meeting of the American Education Research Association. April 8-12, Los Angeles, CA.

Sawaya, S., Bush, J., Jacobs, J. & D'Mello, S.K. (Submitted). Talk Trees: Supporting Growth in Agentic Student Discourse. In Evaluating and Assessing Learning in AI-Embedded Environments: Perspectives across learning scientists. Proceedings of the 20th International Conference of the Learning Sciences.

Trinh, V.A., He, X., and Whitehill, J. (Submitted). LLM revision of named entities from asr transcripts using phonetic and semantic context. International Conference on Acoustics, Speech, and Signal Processing, 2026.