



iSAT Snapshot

NSF AI INSTITUTE FOR STUDENT-AI TEAMING

Winter 2024



iSAT Team Members Set up the Jigsaw Interactive Agent (JIA) Wizard of Oz Studies and Teacher Interviews in the iSAT Lab

2

In Brief...

From the PI

3-4

Jigsaw
Interactive
Agent

5-7

Strand 1
Updates

Strand 2
Updates

8-9

Strand 3
Updates

Institute-wide
Updates

10-12

Member Spotlight

New Quarter,
New Faces!

Publications &
Outreach

In Brief . . .

A quick look at our activities this quarter!

1. iSAT team members, Emily Watts and Rui Zhang, hosted an interactive curiosity station about robots and machine learning techniques at an AI competition and awards event that was sponsored by one of our partner school districts. First, they presented a Sphero programmable robot, which students interacted with to learn fundamental STEM concepts through play-based learning. Second, they showed how machine learning techniques are used to categorize student utterances using labels that they created.
2. iSAT's Executive Director Peter Foltz was interviewed for a Scripps News TV story on the top tech stories of 2023. With the Artificial Intelligence market continuing to boom, Peter gave his expert opinion on how technologies like ChatGPT are just the tip of the iceberg and what we can expect in the coming years.
3. iSAT's Interim Associate Director Thomas Breideband participated in a Student Engagement through Technology Symposium at CU Boulder. He presented on iSAT and how the Institute is using technology to further student learning. The event featured presentations, a panel discussion, and opportunities for hands-on learning.
4. iSAT's work was featured as part of a collection of stories of RPP research and impact, as told by partnerships within NNERPP (National Network of Education Research-Practice Partnerships) from various settings and contexts. iSAT's partnership with iHUB (inquiryHub) was listed as one of six stories of RPP Impact - with an emphasis on our Community Builder Partner (CoBi) and the Learning Futures Workshops (LFW).
5. The AI Steering Committee for K-12 Colorado has been officially launched. Colorado Education Initiative (CEI) and the AI Education Project (aiEDU) are expanding their partnership and previous collaborations to form a new, statewide AI in Colorado Education Steering Committee. iSAT's Co PI Tammy Sumner is a Committee Member, along with iSAT's educator partner Mai Vu.



Interim Associate Director Thomas Breideband participates in a Student Engagement through Technology Symposium at CU Boulder - sharing all about iSAT and allowing the students an opportunity to engage with iSAT's technology.



Co-written by Acting PI Tammy Sumner
& Interim AD Thomas Breideband

Our work together is entering an important and seminal phase: classroom studies of our AI Partners! WooHoo!!!! We have worked hard for over three years to get to this point. First up is CoBi, the Community Builder, which is currently being deployed and studied in numerous classrooms and multiple school districts. These studies will shed light on the degree to which CoBi supports students to develop, enact and enhance their collaborative problem-solving skills and mindsets, which strikes at the core of the iSAT mission and vision. Special thanks to Monica Ko and Bill Penuel for their efforts to recruit additional districts who are ready and eager to participate in these studies!!

At the same time, we are also conducting multiple lab studies. One set of studies is examining in detail the learning processes and learning outcomes of teams using CoBi. These results will complement and extend the findings from our classroom studies. We are also studying an initial version of JIA in the iSAT lab to inform and shape JIA's dialog policy, response generation, and user interface, among other things. Studies of JIA in authentic environments are scheduled for the summer. By the time fall rolls around, we should be in an excellent position to study both AI Partners integrated into the full range of iSAT curriculum in the final year of our performance period. Stay tuned!

Looking outwards, iSAT is continuing its engagement with students, parents, caregivers, teachers, principals, district leaders, and policy makers to figure out how to navigate this new world of AI. For example, we are in the process of inviting youth from a local non-profit community organization to join us for an afternoon of discussion and interaction around AI-enhanced classrooms. This will include curiosity stations related to our work and the opportunity for youth to spend time reflecting about what they learned and sharing out for us what they think about AI and what they hope for the future.

Jigsaw Interactive Agent (JIA)

iSAT's Newest AI Partner for Small Group Collaborations

In 2023, iSAT rolled out its first AI Partner the Community Builder (CoBi) to be tested in classrooms. Now, in our fourth year as an Institute, the iSAT team is excited to introduce a second AI Partner: the Jigsaw Interactive Agent (JIA). Both AI Partners are designed to support student collaboration, but they each focus on different collaborative activities. CoBi focuses on the relationship dimension of collaboration through the lens of community agreements; JIA focuses on so-called Jigsaw activities, providing more interactive feedback to help students engage with each other to share and build on their knowledge and ideas.

What are Jigsaw Activities & What Are JIA's Intended Goals?

Jigsaw activities are a type of collaboration where students first develop initial expertise and then come together to pool their knowledge in an effort to solve more complex problems. Our iSAT curriculum units are all designed to include Jigsaw activities and we started by studying how students were collaborating during these activities. Our analyses revealed that students often did not engage in much discussion and could benefit from group and individual-level support when they engage in real-time activities.

Our intent is for JIA to boost more effective knowledge sharing that can scaffold engagement and lead to increased social cohesion. We will apply Natural Language Processing (NLP) to student conversations to automatically identify instances where support would be helpful (gaps, misconceptions, pauses). This enables JIA to point student groups to relevant resources. This content safety net and explicit positive reinforcement are expected to bolster individual student confidence and, thus, encourage more active participation in small group discussions. More discussion will reinforce knowledge sharing and student learning across

the board, and be beneficial to social cohesion. Members from each of iSAT's research and development strands have been collaborating intensively to develop a dynamic JIA interface and are looking forward to deploying JIA in authentic classroom settings.

An additional long-term goal for JIA is to identify content-appropriate contributions on the part of individual students that may get overlooked by the group as a whole and to encourage exploration of them as a method of explicitly increasing equitable interactions. The NLP components that Strand 1 is developing for this task include Abstract Meaning Representations (AMRs), Dependency Dialogue Acts (DDAs), and our Discourse Policy. The Strand 1 Discourse Policy outlines alternative group discussion states for JIA to detect, each of which would benefit from a different type of JIA response. The plan is to begin with the most obvious signals (no speech, sparse speech, only one person speaking) and then add increasingly sophisticated AI capabilities, AMRs, and DDAs, to the state detection module.

First Version of JIA - Minimum Viable Product

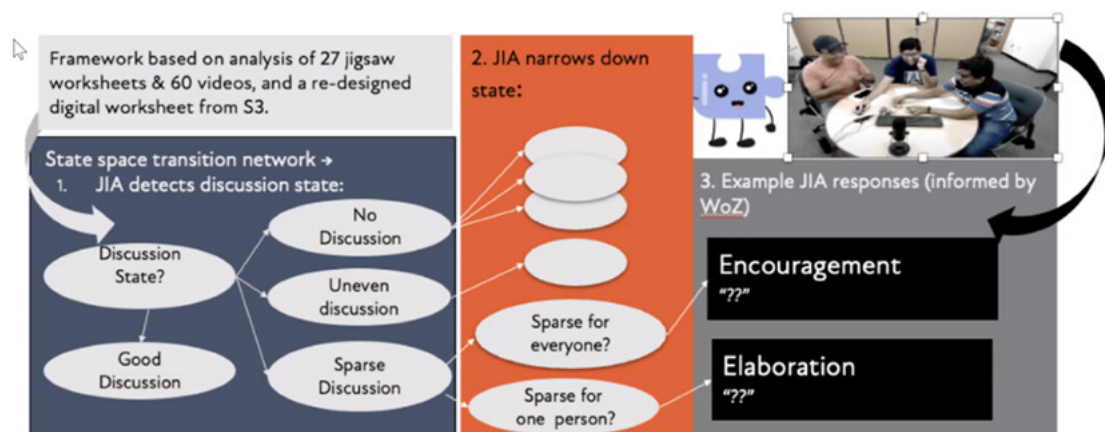
One major area of focus this year has been to build a Minimum Viable Product (MVP). The goal of the MVP is to incorporate our foundational AI approaches into iSAT's existing technical infrastructure and then test and revise them through rapid iterative development. By having this overall infrastructure, new JIA components can quickly be integrated as they are developed. This work has encompassed several steps including:

Identification of team states and feedback. The team has worked to identify how jigsaw groups enter into different states of collaboration and what kinds of feedback would be most beneficial. For example, if JIA sees that the discussion is sparse at a particular point in time, what sorts of statements should JIA make to promote more discussion?

JIA Iterative design process

Left Margin

1) Identification of team states and feedback



The choices for each instance of feedback are derived by reviewing the research literature on supporting collaboration and teacher feedback.

Iterative development of user interfaces. The team iteratively designed a user interface for students that allowed them to work collaboratively on the jigsaw tasks while also interacting with JIA and seeing JIA's feedback. This involved a major revision of the existing paper-based worksheets that students used to fill out as part of these activities.

Wizard of Oz (WOZ). The team developed a WOZ interface where team members could play the role of JIA and interact with students. The interface is being tested in the lab where we can evaluate different forms of feedback and how they are associated with group interactions.

JIA Agent Architecture. Incorporated into our iSAT infrastructure, the JIA agent architecture detects states of collaborations based on variables in our Multimodal Intelligent Analyzer (MMIA). The MMIA analyzes student conversations and generates a number of features such as who is talking, how much they are contributing to the conversation, whether the conversation is on/off topic, and whether collaborative problem solving skills are being exhibited in the discourse.

Discourse Policies and Generated Feedback. With the JIA agent architecture, we are able to integrate discourse policies that analyze the state of student discourse and decide what kinds of feedback should be provided. For example, the discourse policies can detect if a group is off topic or is not advancing new ideas; JIA can decide on feedback that might help the group get back on track. The first versions of JIA incorporate pre-defined feedback while ongoing research developing LLM-based generated feedback will be integrated later.

The expectation is that in a classroom environment, when beginning Lesson 4 of the Sensor Immersion (SI) curriculum, the teacher will explain that there is an AI Partner, JIA, available to help each group with their activity.

What is Next for JIA?

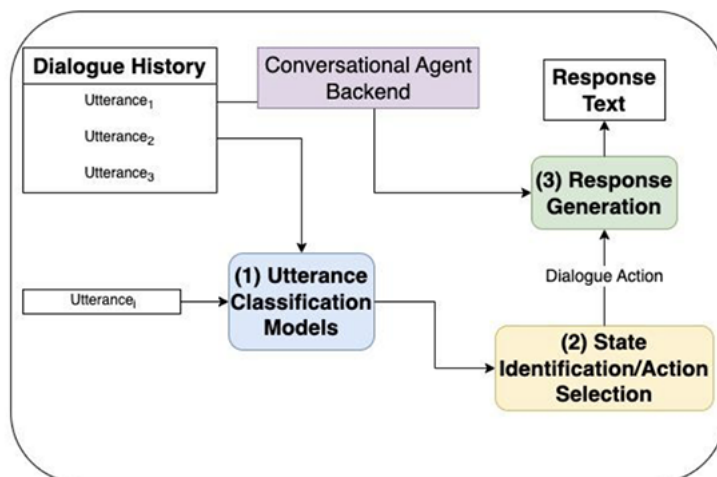
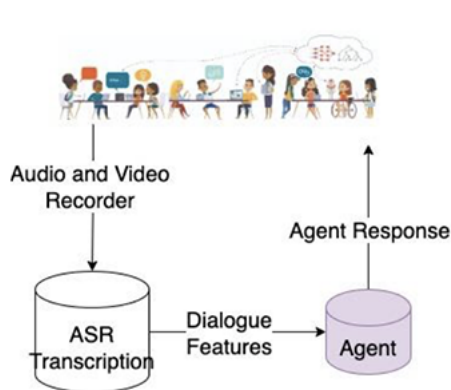
We are continuing to conduct WOZ studies with JIA in the lab with middle school students to refine the development of the user interface, better understand the states of student collaboration, and test different types of feedback. The JIA team is also completing the work of

integrating the discourse policies and feedback into the architecture and linking it with the JIA user interface for the MVP. We will start testing the MVP version of the JIA agent with middle school students in the lab in early spring.

Concurrently, we are conducting a number of evaluations of the individual components that make up JIA as well as JIA as a whole. For examining individual components, this work includes such approaches as 1) evaluating the accuracy on high quality human transcript data of student speech compared to the output of automated speech recognition; 2) measuring the accuracy of automated classifiers such as Collaborative Discourse Moves and On-Task against high quality human annotations for the same tasks; and 3) examining the acceptability of automatically generated JIA responses through human ratings of coherence, naturalness, and helpfulness. To examine the effects of JIA as a whole when students use it, we are evaluating measures like student self-reports of team cohesion, trust of team members, trust of JIA, automated measures of equity in group discussion by examining JIA's effects on turn taking, and the influence or uptake of JIA's feedback by measuring change in types of discourse moves that occur after feedback. Finally, we will be looking at the output of the students completing the tasks, examining their accuracy and completeness and the range and quality of ideas they generate.

JIA in the Classroom

We are currently working with middle school students in the lab and are preparing to transition JIA to classrooms. But first, we are working with iSAT's Teacher Advisory Board to better understand how teachers envision integrating JIA into their classrooms. In these sessions we explore with teachers what kinds of support JIA should provide to student groups and how to present JIA feedback effectively. While JIA has been initially developed with the SI curriculum, we are concurrently adapting it to work for our Games Unit as well. This process is helping the team learn which Jigsaw activities support transfer across tasks and what components may need to be adapted to each curriculum. The team is looking forward to getting JIA into classrooms and seeing how this new AI Partner can be fine-tuned to increase knowledge sharing and team cohesion.



Listen, Understand, and Connect

Strand 1

Strand 1 is guided by the foundational question: What advances in Artificial Intelligence are needed to facilitate collaborative learning conversations? The three research themes identified to help Strand 1 answer these questions are: Speech Processing & Diarization, Content Analysis & Dialog Management, and Situated Grounding.

Speech Processing & Diarization

To understand and support small group collaboration, it is critical to be able to reliably hear what all students in a group are saying. In noisy classroom environments this is a very difficult task! Continuing their work on improving the filtering out of unwanted noise, team members Viet Anh Trinh and Rosy Southwell are developing a SpeechLLM (Large Language Model) that unifies text and speech processing in a single model. This model and training algorithm can benefit from unlabeled speech data using an unsupervised learning approach that is fundamentally not available to standard Automatic Speech Recognition (ASR) architectures such as Whisper.

Research Professor Wayne Ward is exploring how well pre-filtering of speech using a SepFormer, which is a transformer-based neural network for speech separation, can improve downstream ASR accuracy. Tests on the MyST dataset of noise-augmented child speech show promising results. Student Researcher Yiwen Guan is conducting experiments on how to use active learning to accelerate the labeling and training necessary to create ASR systems. It turns out that a random sample of the entire training set is a difficult baseline to beat, but a technique designed to balance how informative versus how representative a particular data sample is, shows some hope of beating it.

Another development this quarter is that Worcester Polytech Institute (WPI) PhD student Xinlu He is exploring new methods for who-is-where-when in group interactions. She is comparing clustering-based methods - which can harness richer geometric structure - to simpler verification-based methods that match each runtime embedding to the set of enrollment embeddings. The team is submitting a manuscript on this work to the Educational Data Mining conference in 2024.

WPI PhD student Jiana Wang is comparing different ways of diarizing an audio transcript of a small-group classroom interaction, with the goal of identifying who-is-speaking-when as well as the aggregate analysis of how-much-did-each-person-speak. The most successful approach diarizes speech at the level of Whisper-generated sentences rather than individual audio frames. The correlation between automatically detected and actual (human-labeled) amounts of speech is 0.65. The team is submitting a manuscript on this work to the Educational Data Mining conference in 2024.

Content Analysis & Dialog Management

The Dependency Dialog Acts (DDA) team has released

the first fully adjudicated DDA annotation set focusing on a physics classroom scenario where students are learning pulley systems. All annotations have undergone thorough discussions and adjudications. The team has reached an edge identification agreement rate of a 0.87 F1 score and a label identification agreement rate of a 0.78 F1 score. This serves as empirical confirmation of the feasibility of DDA even with a relatively large label set and structural information. They are now actively working on developing a large language model (LLM) based parser to address the speakers' intention identification task.

The Abstract Meaning Representation (AMR) team is developing annotations to distill unstructured knowledge resources for educational objectives. They have started by analyzing Sensor Immersion curriculum resources. They are also working towards an AMR verification system to further extend the power of AMR as the critical component of the explainable AI system of iSAT. The team is submitting a manuscript to the Association of Computational Linguistics 2024 conference on the foundational methods of the verification work.

Student Researchers Margaret Perkoff and Emily Doherty have been busy developing and implementing the Jigsaw Interactive Agent's (JIA) Discourse Policy module, which specifies the states to be detected and the subsequent actions entailed by each state. This module will be integrated into the Minimum Viable Product for JIA (MVP - see the JIA feature on page 3), to choose the actions/responses based on the detected states. Margaret, along with student researcher Angela Ramirez, has also been developing a response generation module for JIA, that, given a specific type of interjection, can generate alternative, context-sensitive phrasings. The inclusion of this module is planned for later versions of JIA.

Situated Grounding

Understanding the common ground that student groups establish when interacting with one another, such as their behavioral and verbal cues, as well as prior goals, expectations, and beliefs they bring into the classroom, is the job of the Situated Grounding team.

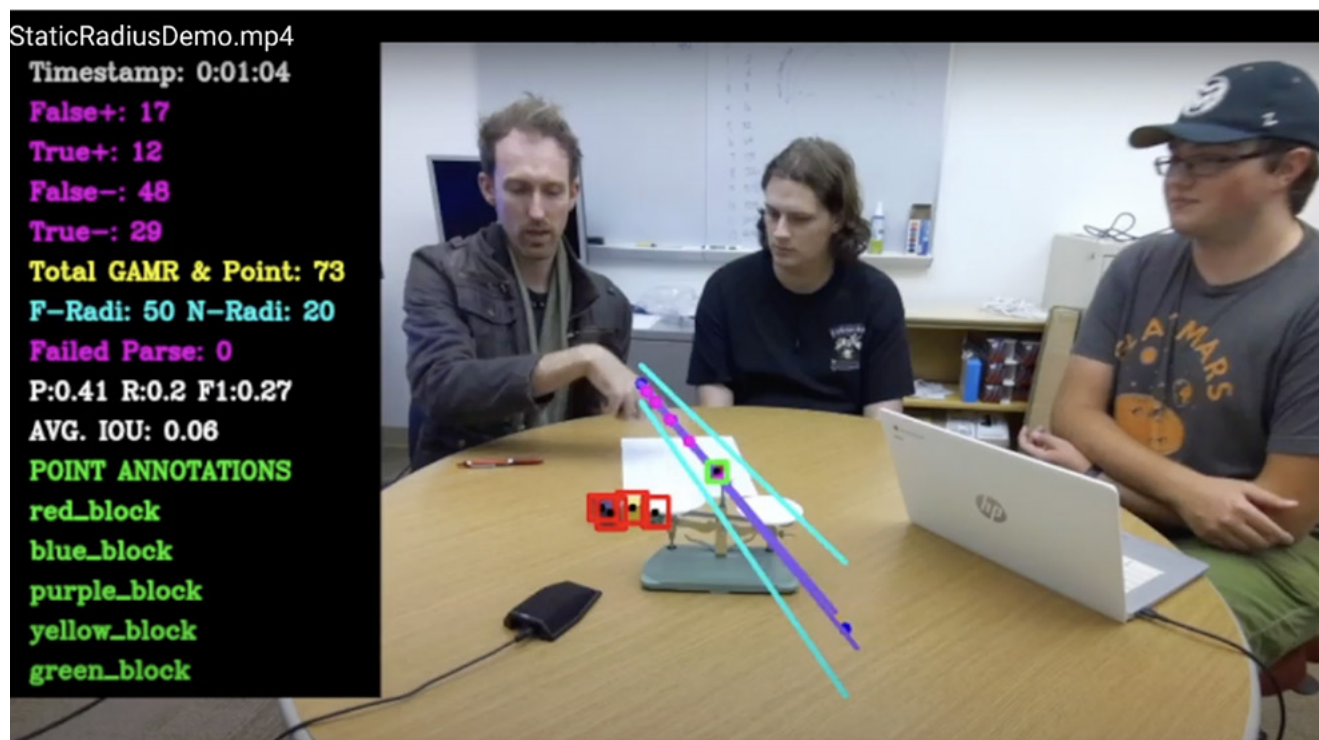
With regard to the multimodal modeling of Common Ground, the team has been working to both model and encode information from multimodal channels in classroom interactions, focusing presently on the Weights Task Interaction Dataset. These include: communicative expressions (speech, gesture), jointly perceived actions, and nonverbal behaviors (gaze, pose, facial expressions). This annotation is being performed in order to train machine learning algorithms to do the following: 1) identify intentions, goals, and attitudes of students, 2) track shared knowledge about tasks and goals, and 3) update evidence and beliefs from actions in context.

Using the information gathered from the above research, they are building a “common ground tracking” algorithm to identify the shared belief space held by all of the participants in a task-oriented dialogue - the task-relevant propositions that all participants accept as true. This consists of: 1) annotating a dataset of multimodal interactions in a shared physical space with speech transcriptions, prosodic features, gestures, actions, and facets of collaboration, and operationalize these features for use in a deep neural model to predict moves toward construction of common ground, 2) model outputs cascade into a set of formal closure rules derived from situated evidence, belief axioms and update operations, and 3) empirically assess the contribution of each feature type toward successful construction of common ground relative to ground truth, establishing a benchmark in this novel, challenging task.

The team has also been working on object selection with pointers. Detecting where students are pointing could be an important indicator of the focus of attention and dialogue in a collaborative task. When aligned with other modal channels like speech, this can help disambiguate referring expressions and demonstratives like “this one”, “that one”, etc. They have developed a method of detecting when subjects make certain gestures, including pointing, and extended that to a method of selecting objects

in 3D space based on the direction of detected points. This includes: 1) detect the joints of participants’ hands and extend a vector through the index finger out into the task space, 2) a frustum is constructed around this vector as the finger moves through space and objects within the frustum region are returned as “selected”, 3) empirically assess the intersection over union of selected objects relative to the human-annotated ground truth, and 4) achieve an IoU value of up to 0.69, though there is considerable variance across groups, due to different pointing strategies in use.

Lastly, in order to improve the content of speech input captured from ASR during the dialogue, the team has employed Dense Paraphrase Enhancement to augment speech input with situated and grounded references from the annotation and redescrptions (paraphrases) of speech transcriptions. This enhances the lexical surface content of the interactions, improving model-based sentence embeddings and subsequent classification of move and engagement types. Examples are as follows with transformations from original transcription (first) to dense paraphrase (second): is [this] one a twenty → is blue block one a twenty, seems like [it] might be about the same → seems like red block, blue block might be about the same, and so [this one] is noticeably heavier than... → so yellow block is noticeably heavier than...



This is a demonstration of the Situated Grounding’s team’s ability to automatically detect when someone is pointing and infer the object that is being pointed at by extending a cone in the direction of the point and seeing which objects fall inside that cone. The selected objects are likely to be the current focus of dialogue.

This can be useful for disambiguating speech or determining the current focus of attention within a collaborative task.

Orchestrating Effective Interactions

Strand 2

Strand 2 is guided by the foundational question: What advances in theories, interaction-paradigms, and frameworks are needed to orchestrate effective student and teacher interactions with AI partners? The three research themes identified to help Strand 2 reach this goal are: Dynamic Framework & Measures of Collaboration, Collaborative Learning (non-verbal and verbal communication; peer scaffolding), and UX Design & Multimodal Modeling.

Dynamic Measures of Collaboration

This team has been working on three main projects. The first project is focused on expanding the influence metric by incorporating non-verbal behavior (eye gaze) and potentially semantic communication analysis (e.g., “uptake” of ideas). To accomplish this, team members have been using iSAT Lab data from the 2023 Data Jam and are developing a multimodal measure of influence, defined as the degree to which individual team members can change patterns at the team level. This work is motivated by the proposition that in social interactions influence is a generalizable coordination mechanism that flows through speech, direction of eye gaze, and gesture. The team has also computed all influence measures for the MakeCode trials in the dataset for influence between participants’ eye gaze (AOI fixations) and communication (presence / absence) patterns. They are currently running statistical analyses to determine whether speech, eye gaze, and/or a combination of speech and eye gaze influence between team members are most predictive of team performance scores and subjective measures of collaboration. The results are being written up for publication in a proposed special issue of Discourse Processes entitled “A Multi-Modal, Multi-Theoretic View of Collaborative Discourse.”

The second project the team is working on incorporates “first-person” (i.e., student) perceptions of collaborative problem solving (CPS) to advance a multilevel perspective of collaborative learning that integrates student perceptions of CPS, objective interactive process measurement during CPS, and existing CPS measurement theory. The project examines a three-level measurement framework that seeks to combine students’ perceptions of collaboration (event cognition), objective interaction processes (speech turn taking), and CPS measurement theory (communication content annotations) to advance a multilevel theory and model of collaborative learning. This project is being led by Julie Harrison for her dissertation. Julie successfully proposed her dissertation in December of 2023. The team has secured IRB approval to run these (Julie’s) studies in the iSAT Lab at CU Boulder. The protocol, stimulus materials, and surveys are in place, with pilots beginning this week. Data collection is expected to be well underway by the end of February.

Collaborative Learning

This theme focuses on identifying verbal and non-verbal modes of collaborative engagement and identifying peer scaffolding moves, as well as understanding how peers support each other during collaborative learning. They are continuing their work on video annotation to help support these goals.

In the iSAT Lab

The multimodal modeling team is preparing a special issue proposal to a high impact journal that takes a Multi-Modal, Multi-Theoretic View of Collaborative Discourse, as mentioned previously. Collaboration is an inherently multifaceted, complex, and multimodal process involving both cognitive and social processes as they occur among multiple people. To accurately investigate and quantify key aspects of collaboration, it is crucial to consider multiple modalities. While speech is an important component of collaborative teamwork, successful collaboration relies on much more than speech and dialogue alone. In addition to their spoken behaviors, humans layer their collaborations using rich non-verbal behaviors such as eye gaze, shared attention, gesture, and facial expression.

The special issue will thus be based around the driving research question on how to measure and model verbal and nonverbal indicators of collaborative problem solving as it unfolds in groups. The team will investigate this question with a series of articles in their special issue that apply different modeling techniques to different modalities (e.g., speech, eye-gaze, nonverbal behavior) collected from the same multimodal collaboration dataset. The co-author teams have worked on the data from the iSAT Data Jam that was held in the spring of 2023 at Arizona State University, applying a variety of modeling and analysis techniques to the rich multimodal dataset.

Strand 2 is also heavily involved in the design of the iSAT Jigsaw Interactive Agent (JIA), with two interconnected areas of focus: 1) running iSAT Lab studies, and analyzing the resulting data in support of the iterative JIA dialog policy creation, and 2) JIA Interface Design. In collaboration with the iSAT infrastructure team, the Strand 2 front-end developer built the JIA web interface where participants in the lab can work with the JIA Partner on SI Jigsaw activities. The JIA web app allows human wizards (played by education experts from Strand 3) to monitor experiment sessions and provide wizard/JIA interventions in real time for the participants. They currently have 8/15 groups (ages 12-17) participating. They are now analyzing the study data (with particular focus on the wizard interventions) in order to iteratively update the JIA dialog policy, with the JIA MVP planned to be rolled out in lab studies at the end of February.

They are also hard at work on the JIA interface design and continue to progress with all participants in the lab studies providing invaluable feedback about JIA through focus groups that are held after they interact with JIA. In addition to the lab studies that provide insights on JIA’s interface and interaction, they are also conducting interviews with members of the Teacher Advisory Board to better understand the strengths and weaknesses of the JIA interface. The interview insights will contribute to the discourse policy and be applied to improve JIA’s interface design.

Engaging Youth

Strand 3

Strand 3 is guided by the foundational question: In what ways can inclusive co-design processes empower stakeholders with diverse identities to envision, co-create, critique, and apply Artificial Intelligent learning technologies for their schools and communities? The three research themes identified to help Strand 3 reach this goal are: Learning Futures Workshop, Games Unit, and Sensor Immersion Unit.

Learning Futures Workshop

After holding three successful Learning Futures Workshops (LFW) every spring over the past three years, the team has recently established a new partnership with a Denver-based youth organization that focuses on community action, healing, and abolition activism. Together with this organization, the LFW team is developing a workshop for Denver youth of color and their adult caregivers, during which they will be exploring how they might build tools that can support community members in collectively managing AI in classroom contexts. This workshop will use home-learning modules (where youth teach AI to their parents) that the Learning Futures Workshop team developed last fall with iSAT interns. In addition to this partnership with the Denver-based youth organization, they have developed a series of *plática* style workshops (i.e., focus groups that center relationship and trust-building through reciprocal sharing between facilitators and participants) with Spanish-speaking parents in the Denver area to explore their concerns and hopes about how AI is used within schools.

Games Unit

The Games Unit team guides the implementation of the Games Unit curriculum - a sequence of nine lessons designed to focus on teaching AI through investigating causes of racism in online gaming communities and what to do about it - as well as holds Professional Learnings for iSAT's partner teachers.

Since December, this team has brought on board two new partner districts to be part of the research on the Games Unit, and both districts will recruit teachers to use the new version and to be working with their first English language arts teacher. They are scheduling professional learning sessions in early spring for one new partner district and a summer workshop for the second. They have also brought on an undergraduate intern to support curriculum revision and professional development activities; this intern's goals are to develop capacities in educational design and research with teachers. In addition to the partnership work with teachers and districts, the team is bringing a jigsaw lesson from the Games Unit into the lab study.

This addition to the lab study is synergistic with the institute-wide effort of developing iSAT's new AI Partner, Jigsaw Interactive Agent (JIA), and will allow the team to test some conjectures about how the task design supports meaningful collaboration. They revised the lesson to fit the lab setting and are modifying the User Interface (UI) so that it can accommodate the lesson. A pilot lab Wizard of Oz (WOZ) study of the Games Unit task is slated for the end of Feb-

ruary. In addition, they have further enhanced the curriculum unit by integrating a new free tool from the MIT media lab that allows students to use block programming to train a classifier to recognize different sentiments. The team is also excited to be piloting the Games Unit for the first time in a school that promotes project-based learning

Lastly, the team, working alongside the Institute-wide group, has developed a CoBi study protocol to evaluate CoBi's efficacy at promoting socio-collaborative learning when piloted in classrooms this spring. Drawing on their expertise in educational validity, they developed, tested and validated measures for this study, then finalized the evaluation plan to be used in both laboratory and classroom settings going forward.

Sensor Immersion

iSAT continues to expand its broader impact by providing hundreds of students with opportunities to learn about AI over the past few months. They had half a dozen teachers implement the Sensor Immersion (SI) curriculum and one teacher implement the Self Driving Car unit, these teachers provided the SI team with continued feedback on curriculum refinement and suggestions for upcoming AI Partner integration. Additionally, iSAT provided over eight hours of professional learning to district partner teachers. During these professional learning workshops teachers and researchers co-designed and refined a new curriculum unit around self-driving cars, developing and refining classroom routines for supporting community agreement establishment and revisiting and supporting implementation of SI. Additionally, they worked with teachers and other members of Strand 3 to refine support for the home learning module, working towards eventual implementation of that module into the Self Driving Cars unit.



Strand 3 held a Teacher Professional Development in January to familiarize iSAT's district partners with the new Sensor Immersion curriculum: Self Driving Cars.

Data Collection, Architecture & Implementation

Institute-wide

The Institute-wide team provides resources (data, annotations, technology) to guide iSAT's core research. The three themes are: Classroom Data Collection and Coordinated Analysis, Technical Architecture, and AI Partner Design, Implementation, and Testing.

Classroom Data Collection and Coordinated Analysis

A constant goal of this team is to maintain a high standard and streamlined data collection process. To help achieve this, data collectors, who have already been trained in classroom data collection and taking field notes, were additionally trained in lab data collection and running lab studies. This enables the team to run studies for assessing the AI Partners in a more controlled environment while simultaneously continuing to collect data from live classrooms. The data collectors also reviewed procedures for testing different microphones during classroom data collection that will occur in the spring. Improvements made to the transcription process last fall allowed for development of a new standard of transcription, called MidFi, that contains enough information for data annotation and analysis in addition to speeding up the process for creating the HiFi transcripts necessary for automatic speech recognition (ASR) testing.

For improving and speeding up the Collaborative Problem Solving (CPS) annotation process, more annotators have been training with this method. Previous applications of CPS annotations have focused on one lesson or type of CPS activity at a time. In preparation for spring 2024 data, the team is extending their annotations to include a broader range of lessons and activities. A review of research operationalizing CPS measurement through human coding of small group verbal discourse continues work on CPS framework development, which will culminate in a manuscript demonstrating a computational method for bringing together multiple approaches to measurement of a single construct.

Technical Architecture

The technical architecture team has made quite a few key advancements over the past several months. They have made additions to the Researcher Data Display for the AI Partner Community Builder (CoBi), which include the following: 1) a visualization tool that shows the activity of various speakers throughout a recording session, aiding researchers in understanding speaker interactions, 2) a new sentence cloud visualization within the Researcher Display and CoBi interfaces, highlighting CoBi noticings for comprehensive analysis, 3) a display for the Abstract Meaning Representation (AMR) Analysis that offers deeper insights into semantic structures, and 4) a pie-chart placeholder for visualizing influence data, enhancing the display capabilities of research findings.

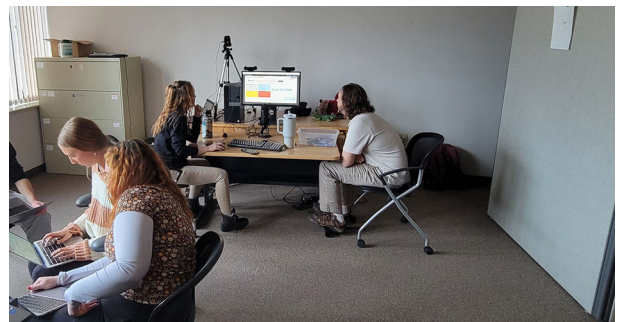
The Real-Time Analysis Pipeline for CoBi also has several improvements in the works to bolster real-time audio and video processing capabilities in lab and classroom environments including developing a real-time audio transcription processor compatible with Mac systems - facilitating seamless testing procedures. They are also currently integrating the

Multimodal Intelligent Analyzer (MMIA) pipeline with Faster Whisper processor for GPU-optimized audio transcription. Performance metrics will be collected to assess processing efficiency and agent response times. Regarding AMR integration, the team has been working on establishing infrastructure for AMR analysis of student interactions and integration with the MMIA analysis database. They have also integrated AMR data into the Research Data Display, enhancing the depth of analysis available to researchers. Currently, and over the next several months, they are developing a visualization tool to display AMR analysis results within the Researcher Data Display.

AI Partner Design, Implementation & Testing

This team has recently rolled out an improved version of CoBi last fall. Currently, they are implementing CoBi Classroom Studies, which is a quasi-experimental study in middle school classrooms with some teachers teaching the Sensor Immersion (SI) curriculum unit with the CoBi AI Partner, while other teachers implement the curriculum unit without AI support. This is meant to provide valuable data on potential learning outcomes and identify areas to improve CoBi further. Alongside the classroom study, they are conducting a CoBi efficacy lab study with undergraduate students as well as high school and middle school students. This study follows a random assignment design with two control groups. One control group is introduced to community agreements but receives no feedback. The other control group receives feedback through CoLe (short for Communication Level) – an alternative form of collaboration feedback based on indicators including verbosity, speaking time per speaker, and speaker-switching dynamics. The interface for this alternative feedback is realized within the existing CoBi web application.

In addition to the CoBi studies, the Institute-wide team has been hard at work on the Jigsaw Interactive Agent (JIA) Infrastructure, integrating it with MMIA. This consists of establishing the cloud infrastructure for the JIA AI Partner, covering front-end web applications, DynamoDB databases, and analysis agents. It also implements event-based messaging to relay MMIA analysis results from student audio to the JIA agent, enabling real-time response generation for students via the front-end application.



Institute-wide team members set up the CoBi Efficacy Study.

Learn more about our members!

M eet Azel Reitzig - Executive Director of Innovation at the Innovation Center of the St. Vrain Valley Schools (SVVSD).



Q: How did you get involved with iSAT and in what capacity do you work with the Institute?

A: I was introduced to Tammy Sumner in the fall of 2019. We had been doing work at the Innovation Center around AI. Tammy and I talked and, well, here we are! At first, I was the main point of contact in SVVSD. I recruited teachers and coordinated with the iSAT team on trainings and sessions at schools. Most recently, I have been working with Emily Watts developing the high school intern project, which we are very excited about!

Q: What is your background?

A: I am a Colorado native, born in Denver, and I have 25+ years of K-12 education experience! I earned a masters degree in Comparative Literature from CU Boulder as well as my secondary teaching license in Language Arts and German. I also have a masters in educational leadership

from the University of Northern Colorado.

Q: What do you enjoy most about your current position?

A: The energy at the Innovation Center is very different from other places I've worked. It's very positive and purposeful. These are paid opportunities in which our student designers are doing authentic work for real clients. The students get to test-drive a career and be paid for it. They lead the work with support from our program managers and mentors. It's so much fun!

We also get to be on the bleeding edge of where education, technology and society intersect. One of our primary responsibilities is to explore emerging technologies and see how they impact or fit into K-12 education. Our work with iSAT is a perfect example of this: how might AI allow us to innovate education and help our students be more successful? As we do this exploration, we also are reflecting on how we might scale what works across our system to advance the entire system. It's very real and relevant.

Q: What do you like to do outside of work?

A: I enjoy working and being in my garden. I play soccer regularly with a group of colleagues and friends. I attempt golf. I spend time with my family, cook a lot, which means I eat a lot.

M eet Marlene Palomar - a Strand 3

Graduate Student Researcher who is advised by Dr. Kalonji Nzinga & Dr. Ben Kirshner. In her research, Marlene explores storytelling, creative expression, and healing within learning environments through critical theories and interdisciplinary methods.



Q: What are you working on?

A: I'm currently collaborating on a project with Dr. Michael Alan Chang that explores young people's views for AI in the classroom and beyond. This collaboration is focused on adolescents and their families/caregivers, we will be

hosting workshops conducted by the local youth to better understand how they want AI to be utilized in their schools.

Q: How does your research contribute to iSAT?

A: This research contributes to the interdisciplinary commitment and community dedication by incorporating input from families and caregivers. Previously, we have collaborated with diverse groups of students and teachers to investigate collaborative learning environments, but we are now extending our focus beyond the classroom to encompass the dynamics of the home environment and incorporate valuable insights from families and caregivers.

Q: What is a fun fact about you?

A: I have a deep appreciation for unique jewelry, whenever I travel I get jewelry as a souvenir!

New Quarter, New Faces!

Growing our team and our impact.

We're excited to welcome new members to our team! Here are four featured below.

Brooklyn Clines is a temporary Research Assistant and pursues a passion for the classroom by supporting teachers and researchers through Data Collection. She's also working in teacher outreach and a liaison for the Teacher Advisory Board. Recently, she graduated from the University of Oklahoma, Norman with a B.S. in World Language Education for French. Brooklyn has a love for teaching and second language acquisition.

Shriprajwal Krishnamurthy is a Software Engineer working with the Institute-wide team to improve the backend of iSAT. His focus is on content analysis and dialogue management and ensuring a smooth experience for the users of iSAT's AI partners. He is currently pursuing a Masters in Computer Science at CU Boulder.

Tilak Singh is a first-year research-based graduate student at CU Boulder who is part of iSAT's Institute-wide team as a software engineer. Originally from Lucknow,

India, Tilak completed his Bachelor's degree in Information Science from MS Ramaiah Institute of Technology in Bangalore and worked as a full stack software developer at a health tech startup. His interests primarily revolve around cloud computing and software development.

Sachin Rathod is a Software Engineer working on full-stack development of iSAT's AI Partners with the Institute-wide team. A graduate student pursuing a professional-based master's degree in computer science at the University of Colorado Boulder, Sachin received his undergraduate degree from the National Institute of Technology Karnataka, Surathkal, in information technology in 2019. After his undergraduate studies, he worked as a Software Engineer at SAP Labs India and a Senior Software Engineer at Optum. His interest lies in Machine Learning, Distributed Systems and Cloud Computing.

Welcome Brooklyn, Shriprajwal, Tilak, Sachin!



Brooklyn Clines
Institute-wide
Research Assistant



Shriprajwal Krishnamurthy
Strand 1
Software Engineer



Tilak Singh
Institute-wide
Software Engineer



Sachin Rathod
Institute-wide
Software Engineer

Get
Involved!

Email, subscribe, and follow us to learn more!



info.ai-institute@colorado.edu



www.colorado.edu/research/ai-institute/



@NSF_iSAT

iSAT Publications and Outreach

Published / Accepted

Diaz-Asper, C., Hauglid, M. K., Chandler, C., Cohen, A. S., Foltz, P. W., & Elvevåg, B. A framework for language technologies in behavioral research and clinical applications: Ethical Challenges, Implications and Solutions. *American Psychologist*. 79 (1), 79–91 (2024).

Foltz, P. W. & Chandler, C. (in press). Computational language as a window into cognitive processing. In M. D. Shermis and J. Wilson (Eds.) *The Routledge International Handbook of Automated Essay Evaluation*. Routledge (2023).

Gorman, J.C & Wiltshire, T.J. A typology for the application of team coordination dynamics across increasing levels of dynamic complexity. (At the Forefront of Human Factors & Ergonomics invited mini-review). *Human Factors*, 66, 5-16. (2024)

Harrison, J.L., Gorman, J.C., Reitman, J.G., & D'Mello, S.K. Toward the validation of a novel measure of individuals' influence during team collaborations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. (2023).

Hirshfield, L. AI-Partners to Support Group Collaboration in Classrooms. Accepted in AERA symposium 'The Past, Present, and Future of Intelligent Pedagogical Agents: From Agents to Human Digital Twins. Philadelphia, PA. April (2024).

Lodzickowski, K., Foltz, P. W. & Behrens, J. T. (in press). Generative AI and its educational implications. In T. Kourkoulou, D. O. Tzirides, B. Cope, and M. Kalantzis (Eds.) *Trust and Inclusion in AI-Mediated Education: Where Human Learning Meets Learning Machines*, Springer (2024).

Perkoff, M., Doherty, E., Bush, J., & Hirshfield, L. Crafting a Responsible Dialog System for Collaborative Learning Environments. In *AI for Education: Bridging Innovation and Responsibility at the 38th AAAI Annual Conference on AI*. (2023).

Submitted / Under Review

Chang, M. A., Wong, R., Breideband, T., Philip, T. M., McKoy, A., Cortez, A., and D'Mello, S. Co-design Partners as Transformative Learners: Imagining Ideal Technology for Schools by Centering Speculative Relationships. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*.

Chandler, C., Ganesh, A., Kann, K., Palmer, M., D'Mello, S. Prompting as Panacea? A Case Study of In-Context Learning Performance for Qualitative Coding of Classroom Dialog In *Proceedings of the Educational Data Mining conference (EDM 2024)*.

He, X., Trinh, V. A., McReynolds A., and Whitehill, J. Person Re-identification in Classroom. In *Proceedings of the Educational Data Mining conference (EDM 2024)*.

Perkoff, M. E., A. Ramirez, Doherty, E., and Martin, J. LLM Guardrails: Constraining Generative Output to Create Smarter and Safer Dialog Systems. In *Proceedings of the Association for Computational Linguistics conference (ACL 2024)*.

Wang, J., Whitehill, J., and Dudy, S. (2024) Speaker Diarization for Classroom Group Discussions. In *Proceedings of the Educational Data Mining conference (EDM 2024)*.

Southwell, R., Ward, W., Anh Trinh, V., Clevenger, Ch., Clevenger, C., Watts, E., Reitman, J., D'Mello, S., and Whitehill, J. "Automatic speech recognition tuned for child speech in the classroom. *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (2024).