The Role of Visual Analytics in the New Era of AI-assisted Education

Meng Xia
July 22, 2024

ChinaVis 2024

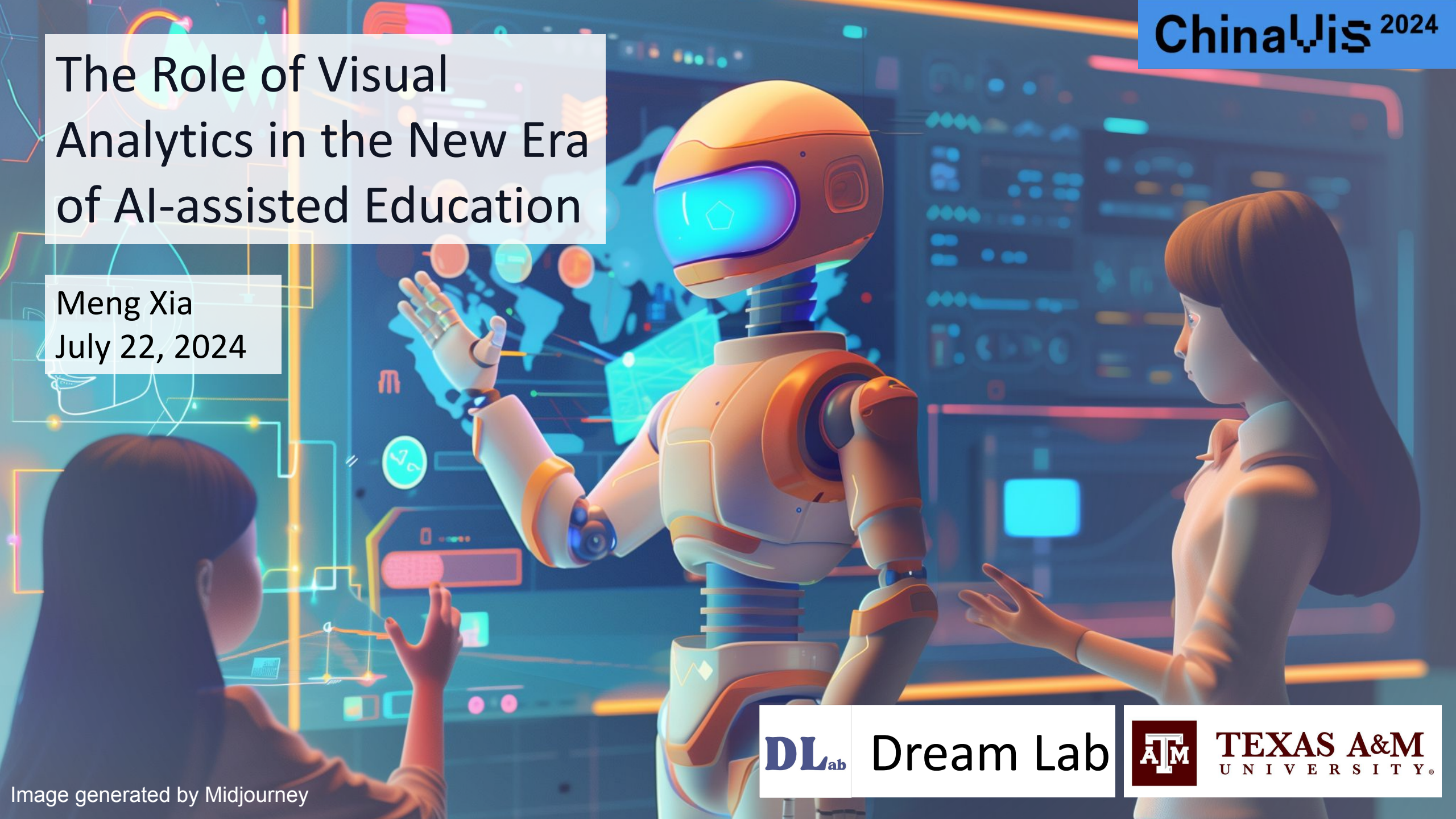
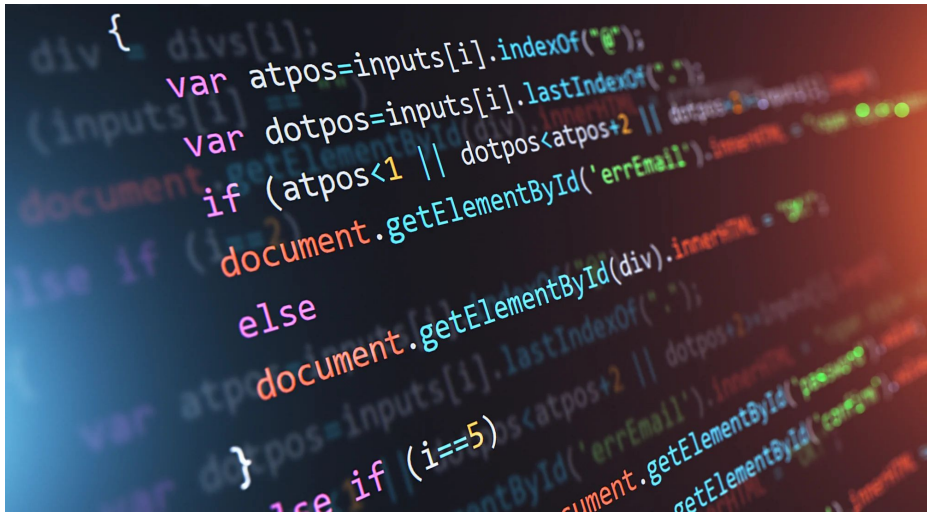


Image generated by Midjourney

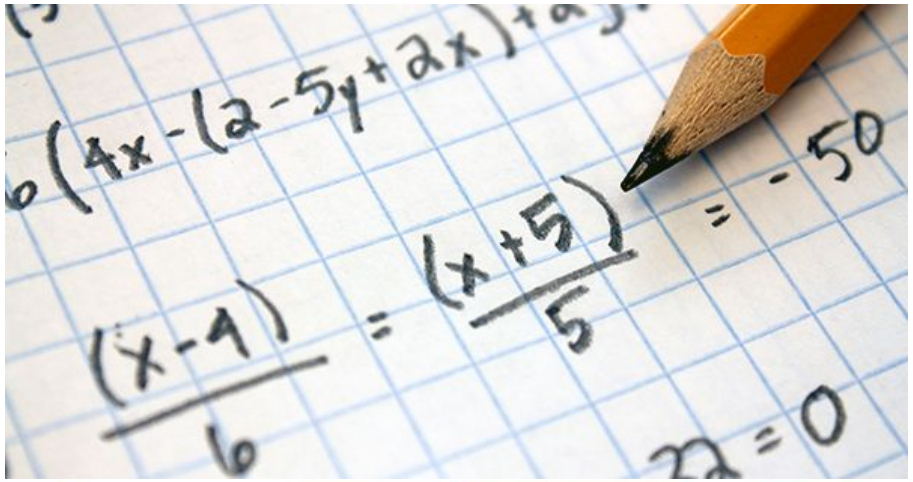
DLab Dream Lab

TEXAS A&M UNIVERSITY



ChatGPT

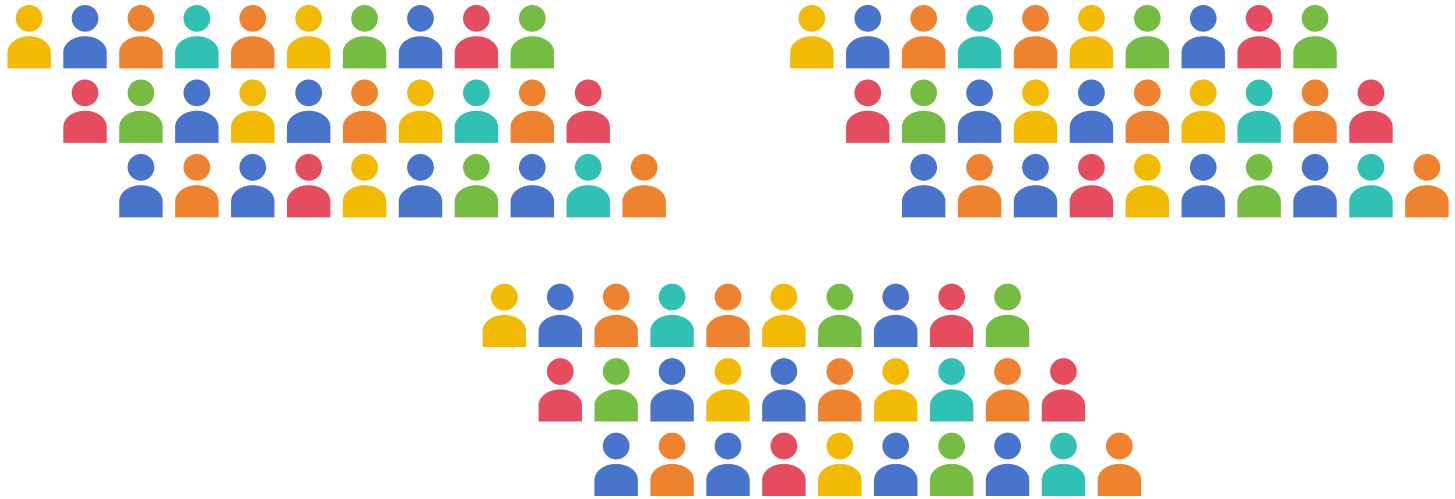
Midjourney



Usage of AI in Education

Education Conference	**Use of AI (Full paper)**	**Use of LLM (Full paper)**
AI in Education 2024 (AIED)	37/49 (76%)	22/37 (60%)
Learning@Scale 2024 (L@S)	16/22 (73%)	12/16 (75%)

Personalization as a Foundational Education Challenge



Why personalization?

Non-cognitive

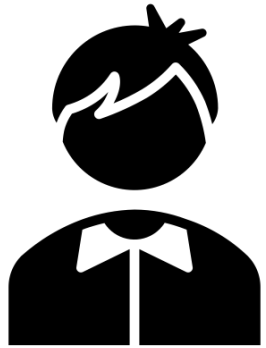
Motivation
(D'Mello, Lehman, Pekrun, & Graesser, 2014)

Self-regulation skills
(Aleven & Koedinger, 2002)

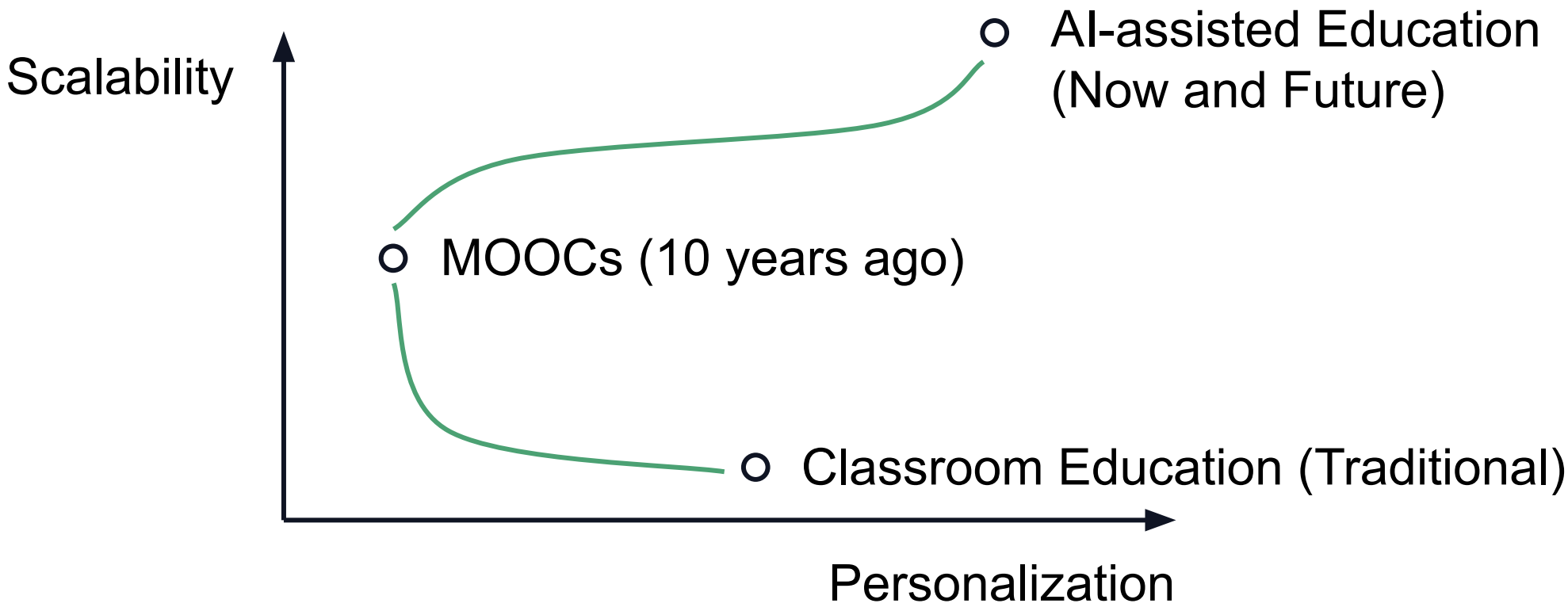
Cognitive

Knowledge
(Koedinger,. Stamper, McLaughlin, & Nixon, 2013;)

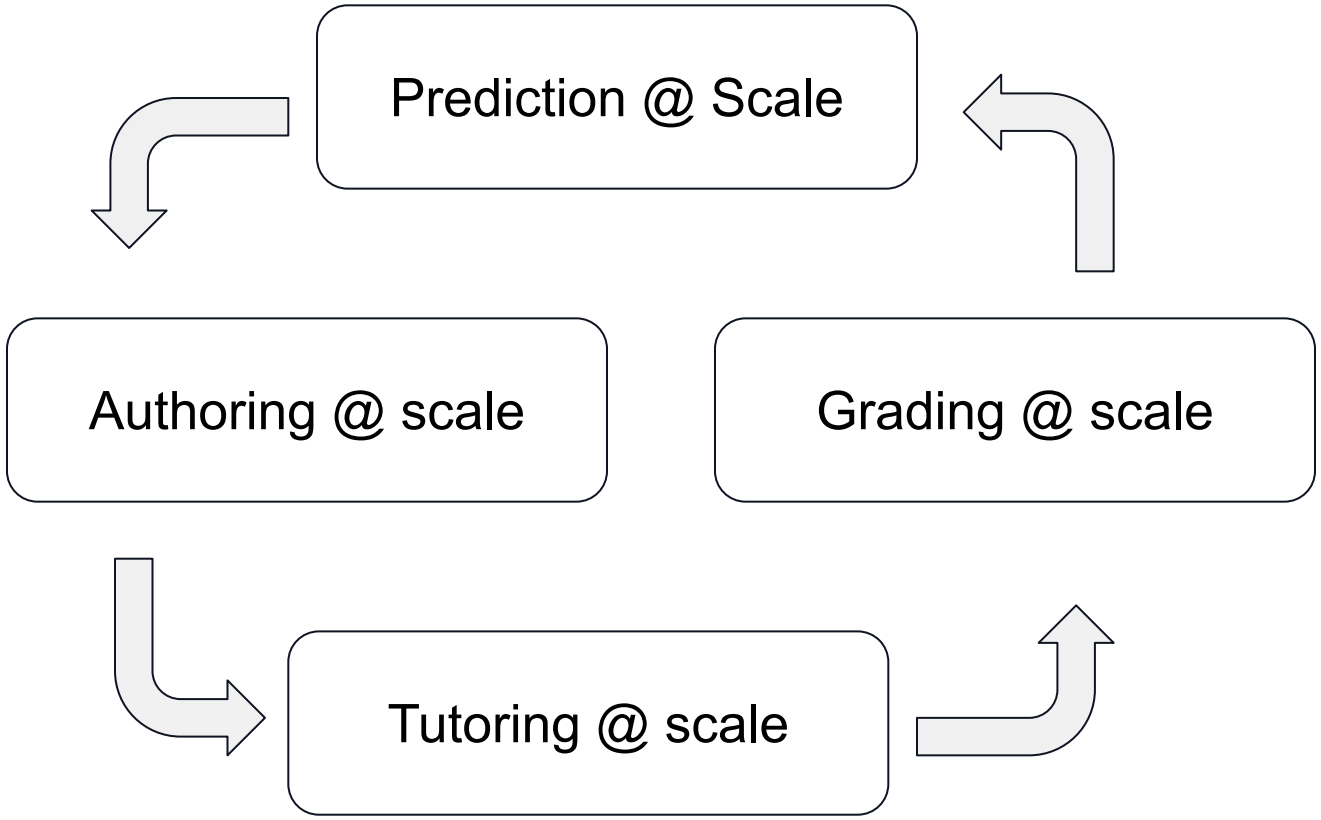
Problem-solving strategies, errors
(Adams et al., 2014)



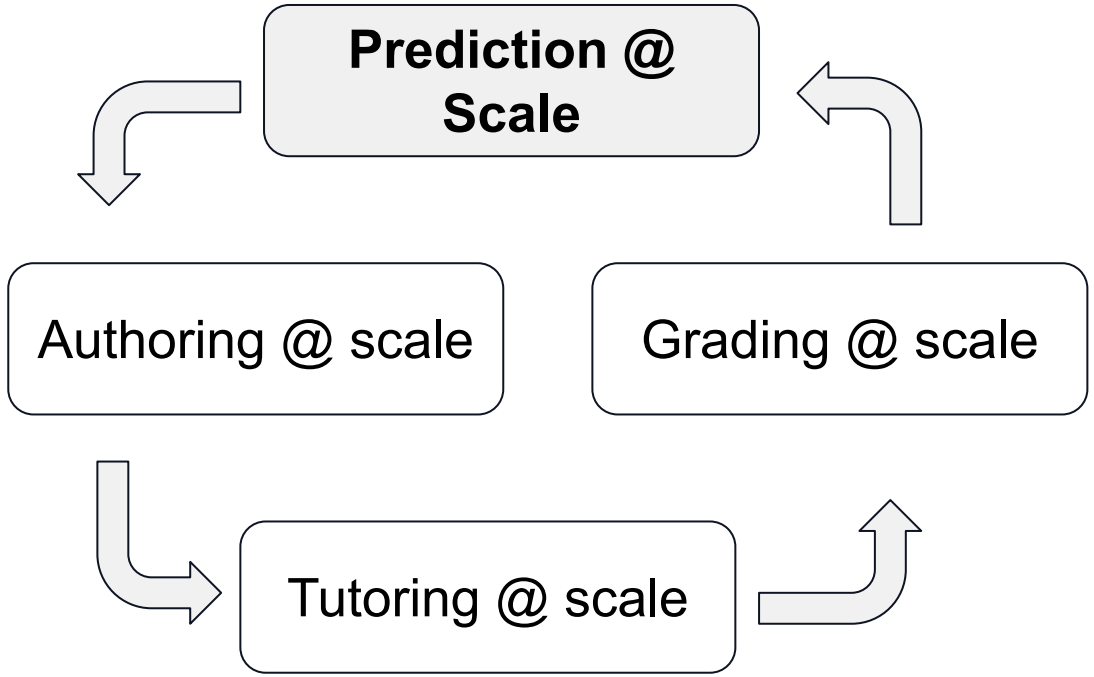
Vision for education: **Personalization @ Scale**

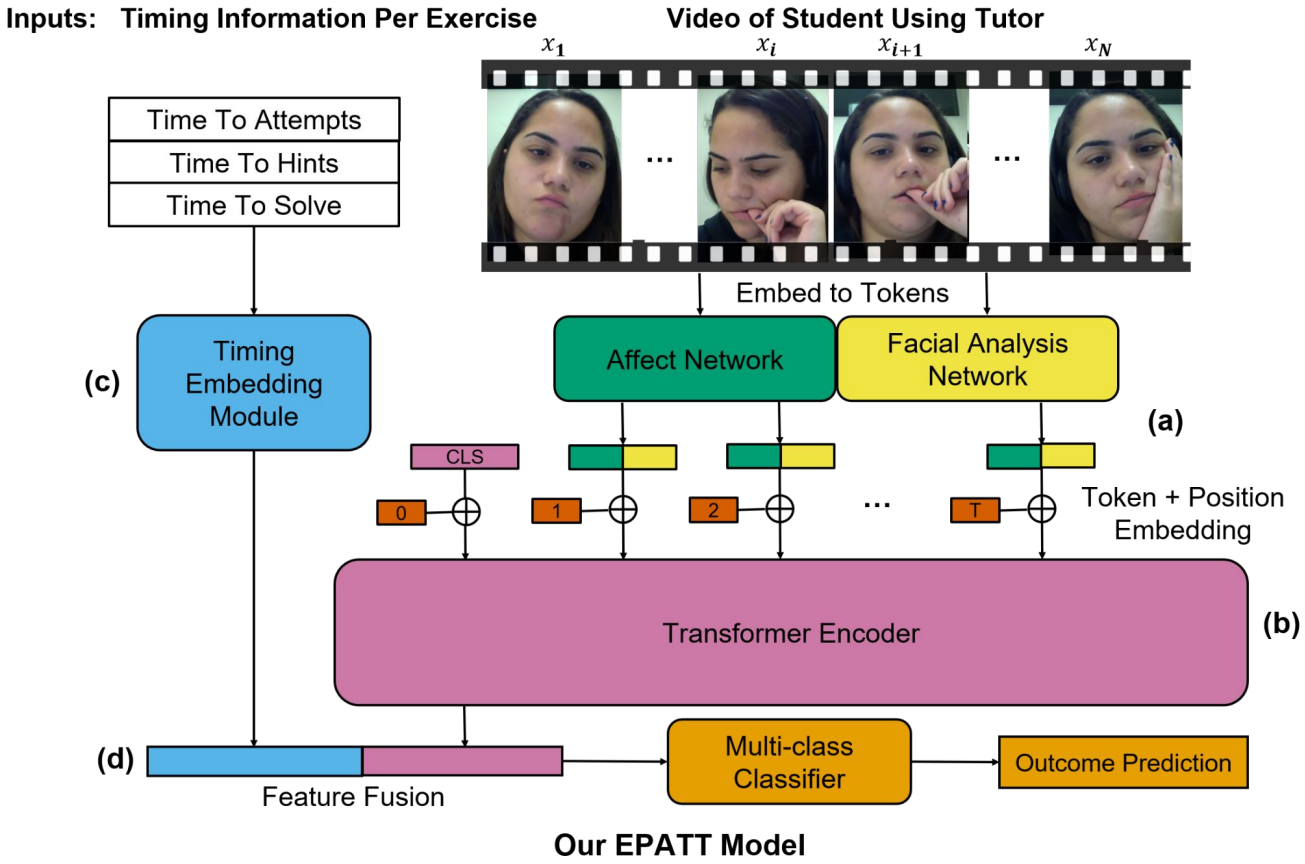


What are the AI's roles?



What are the AI's roles?





Yu, Hao, et al. "Affect Behavior Prediction: Using Transformers and Timing Information to Make Early Predictions of Student Exercise Outcome." *International Conference on Artificial Intelligence in Education*. Cham: Springer Nature Switzerland, 2024.

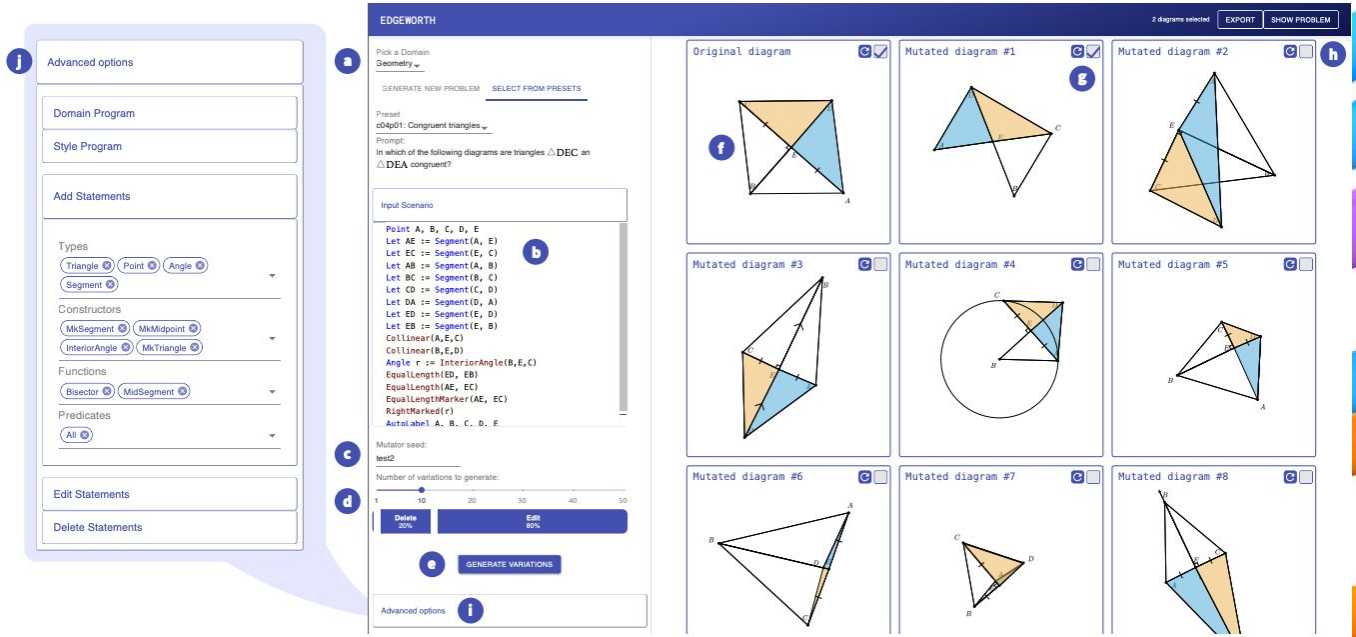
What are the AI's roles?

Prediction @ Scale

Authoring @ scale

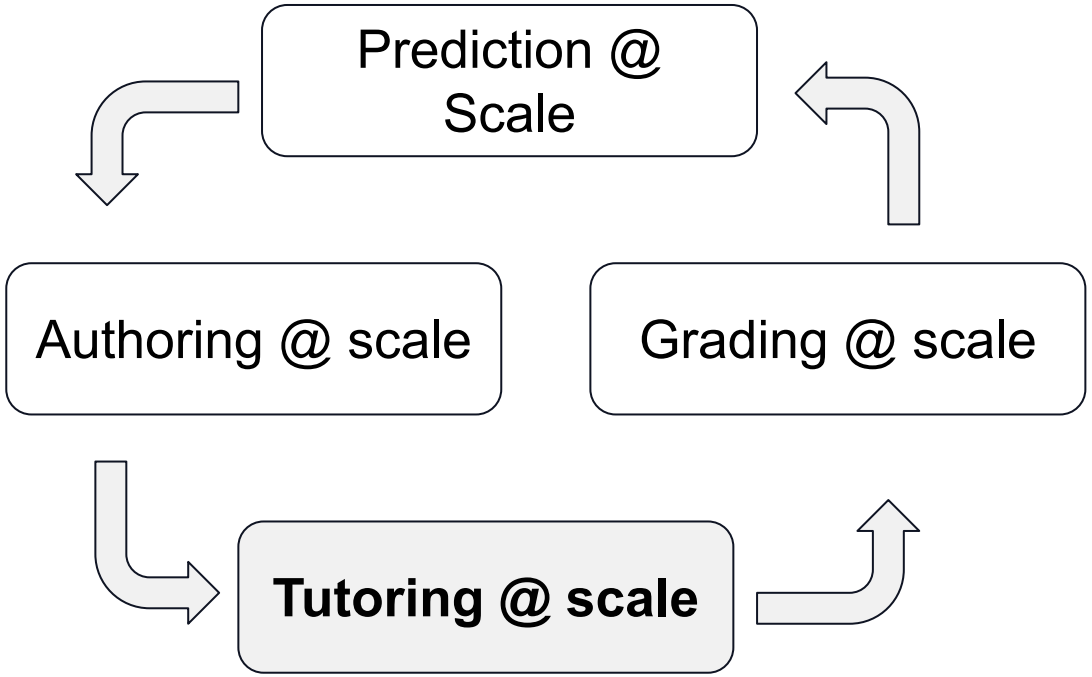
Grading @ scale

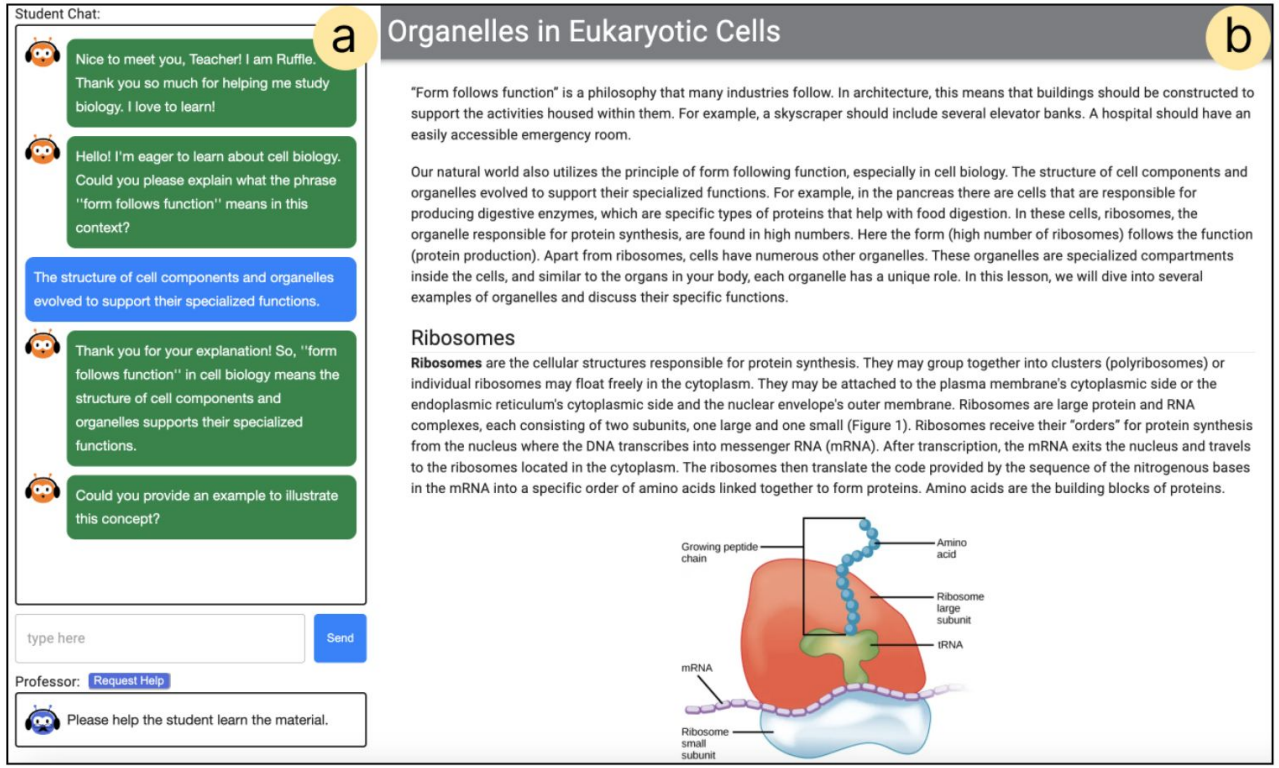
Tutoring @ scale



Ni, Wode, et al. "Edgeworth: Efficient and Scalable Authoring of Visual Thinking Activities." Proceedings of the Eleventh ACM Conference on Learning@ Scale. 2024.

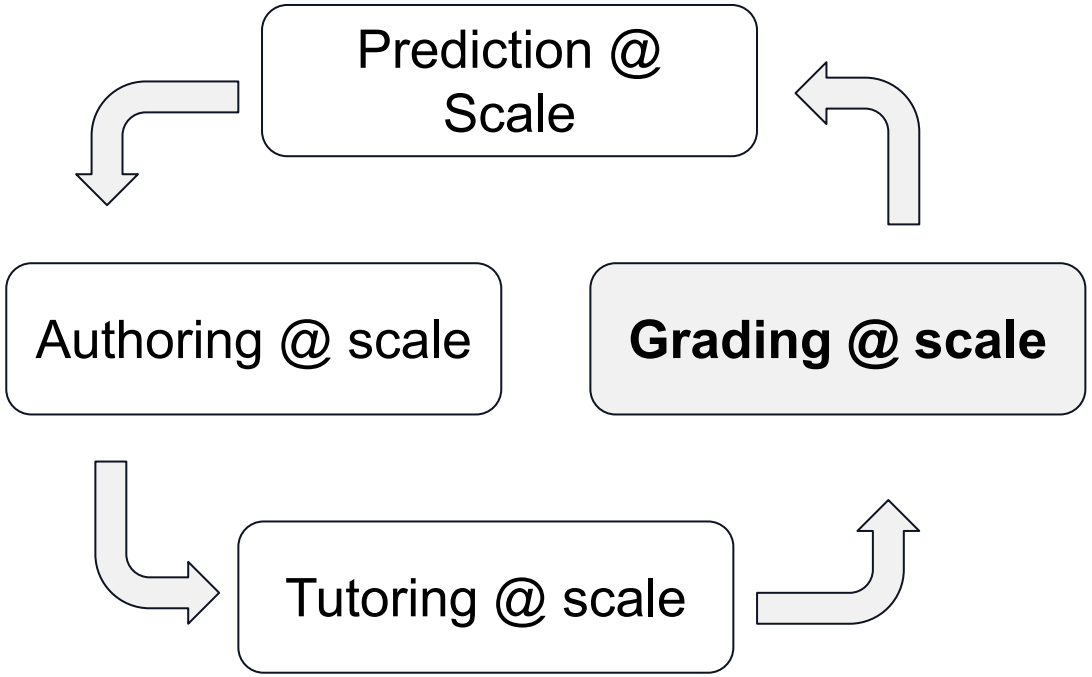
What are the AI's roles?

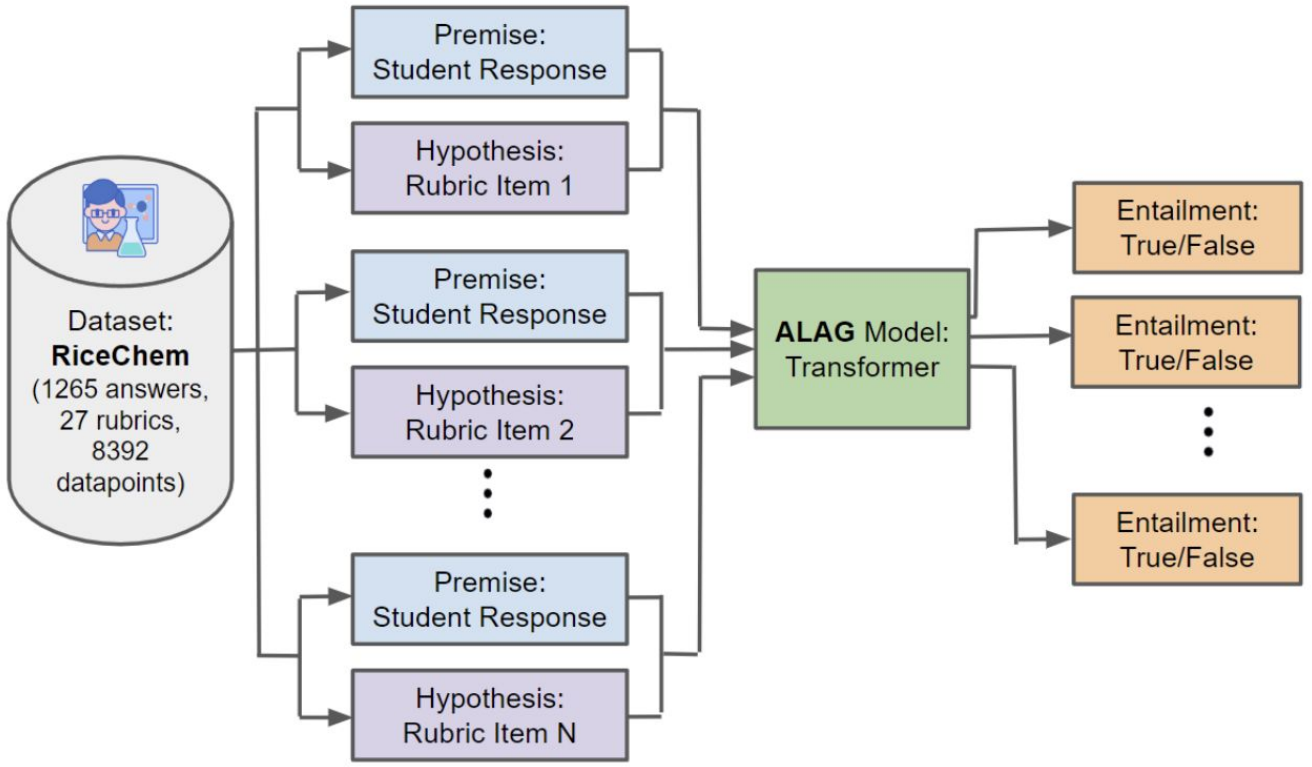




Schmucker, Robin, et al. "Ruffle &Riley: Insights from Designing and Evaluating a Large Language Model-Based Conversational Tutoring System." International Conference on Artificial Intelligence in Education. Cham: Springer Nature Switzerland, 2024.

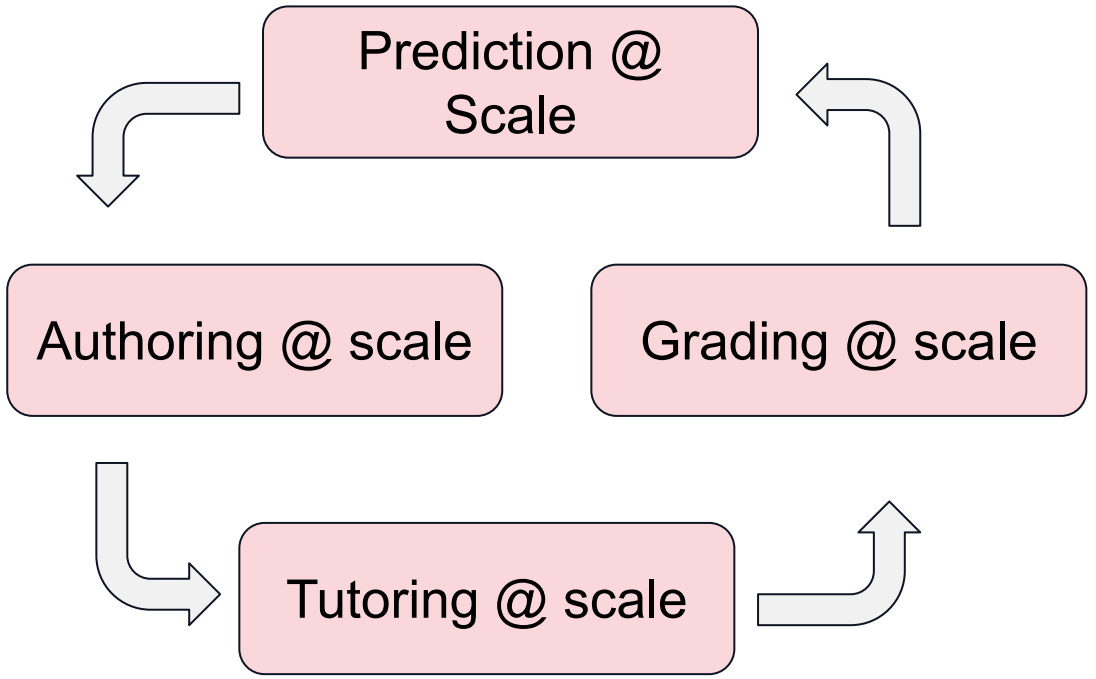
What are the AI's roles?



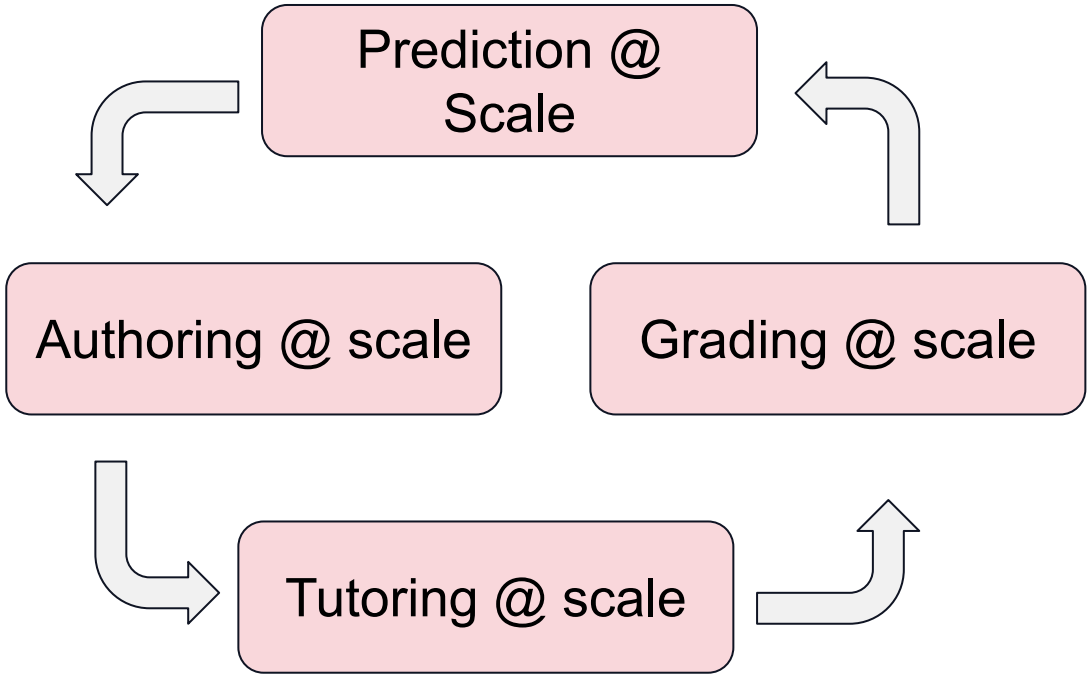


Sonkar, Shashank, et al. "Automated Long Answer Grading with RiceChem Dataset." International Conference on Artificial Intelligence in Education. Cham: Springer Nature Switzerland, 2024.

What are the AI's problems?



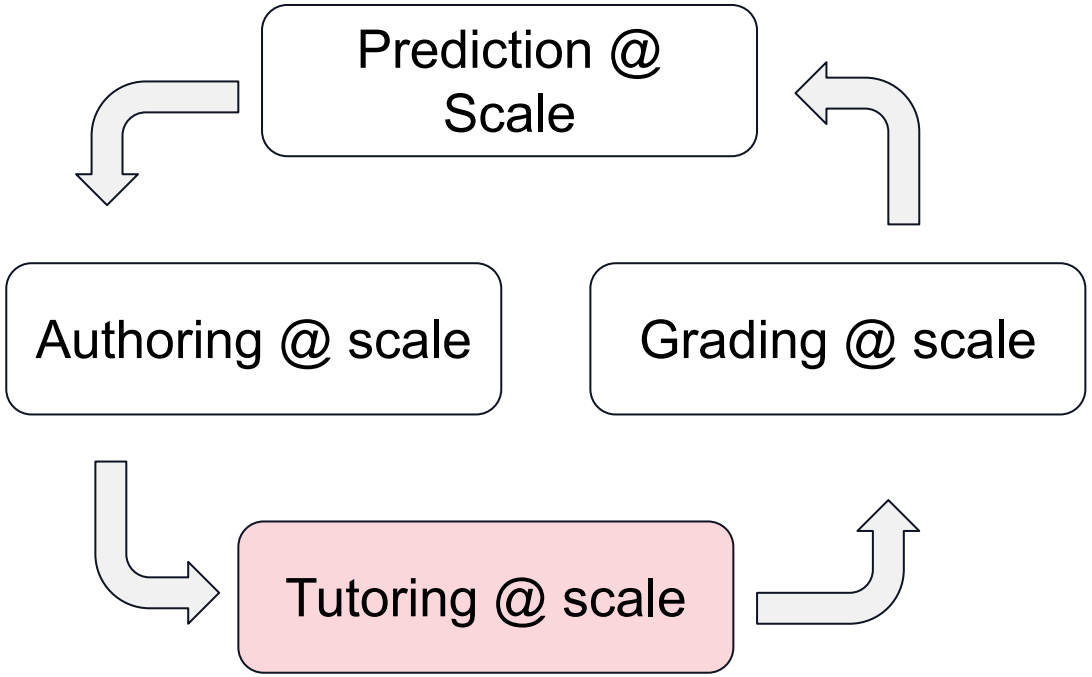
What are the AI's problems?



Content inaccuracy
Improper use of AI

- Cain, William. "Prompting change: exploring prompt engineering in large language model AI and its potential to transform education." *TechTrends* 68.1 (2024)
- Rasul, Tareq, et al. "The role of ChatGPT in higher education: Benefits, challenges, and future research directions." Journal of Applied Learning and Teaching 6.1 (2023): 41-56.

What are the AI's problems?

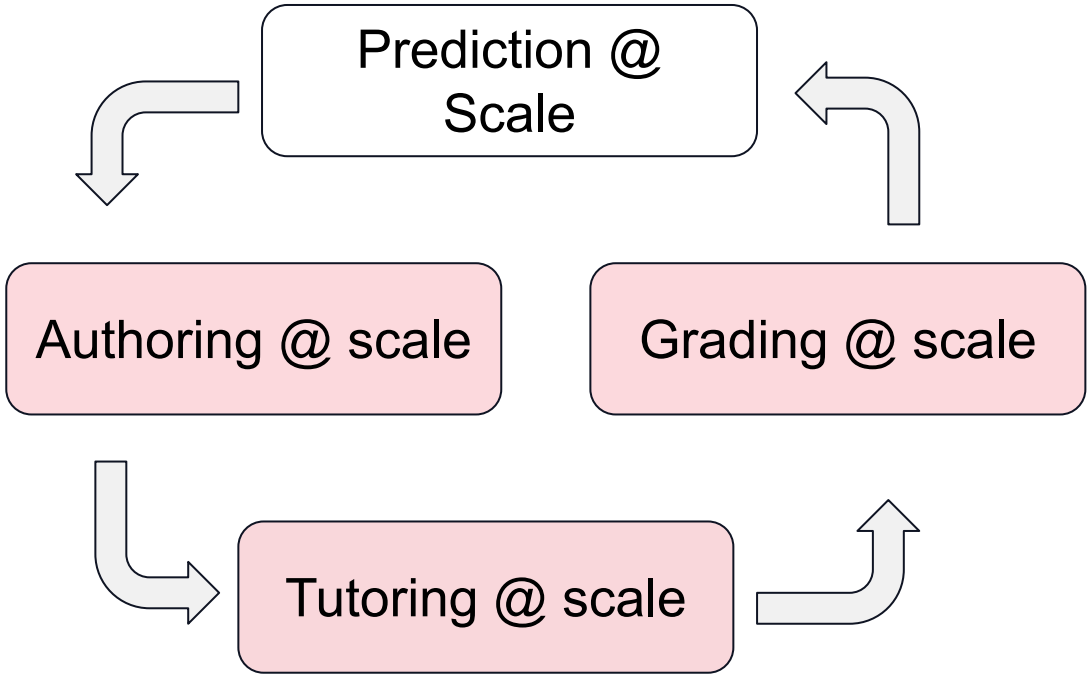


Content inaccuracy
Improper use of AI

Lack of pedagogical guidance
Not personalized

- Cain, William. "Prompting change: exploring prompt engineering in large language model AI and its potential to transform education." *TechTrends* 68.1 (2024)
- Rasul, Tareq, et al. "The role of ChatGPT in higher education: Benefits, challenges, and future research directions." Journal of Applied Learning and Teaching 6.1 (2023): 41-56.

What are the AI's problems?



Content inaccuracy
Improper use of AI

Lack of pedagogical guidance
Not personalized

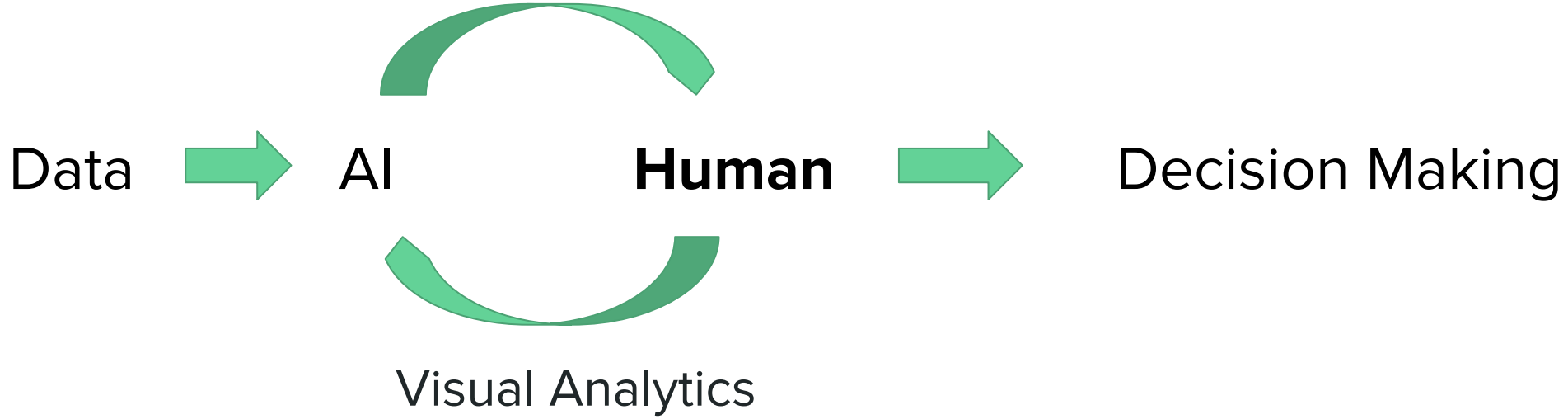
Lack of evaluation
High risk

- Cain, William. "Prompting change: exploring prompt engineering in large language model AI and its potential to transform education." *TechTrends* 68.1 (2024)
- Rasul, Tareq, et al. "The role of ChatGPT in higher education: Benefits, challenges, and future research directions." Journal of Applied Learning and Teaching 6.1 (2023): 41-56.

The role of Visual Analytics: Augmenting Intelligence

Reckoning
(calculative predictions)

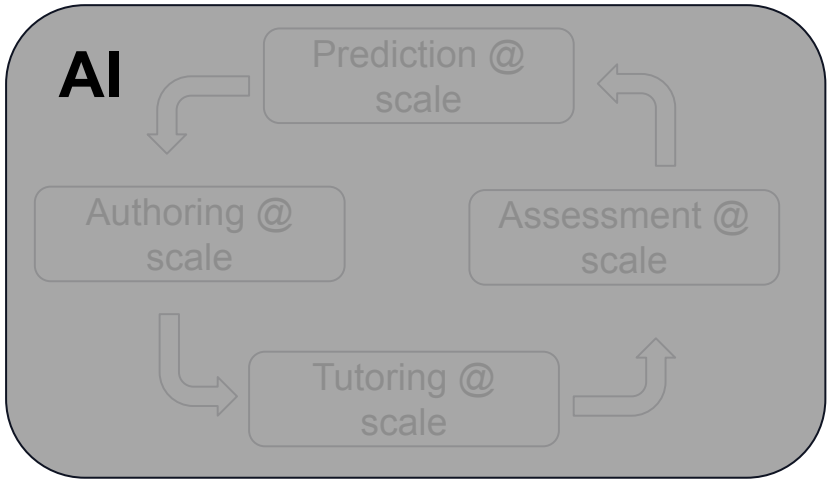
Judgement
(practical wisdom)



Human

Content inaccuracy
Improper use of AI

Analyzing @ scale



Lack of pedagogical guidance
Not personalized

Explaining @ scale

Lack of evaluation
High risk

Testing @ scale

Human

Content inaccuracy
Improper use of AI

Analyzing @ scale

AI

Prediction @ scale

Authoring @ scale

Assessment @ scale

Tutoring @ scale

Lack of pedagogical guidance
Not personalized

Explaining @ scale

Lack of evaluation
High risk

Testing @ scale

AI's role: tutoring @ scale

Vis' role: analyzing @ scale

StuGPTViz: A Visual Analytics Approach to Understand Student-ChatGPT Interactions

Zixin Chen, Jiachen Wang, Meng Xia, Kento Shigyo, Dingdong Liu, Rong Zhang, Huamin Qu

VIS 2024

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Background: An inevitable trend in using LLMs

ChatGPT

Jennifer L. Steele, To GPT or not GPT? Empowering our students to learn with AI, Computers and Education: Artificial Intelligence, Volume 5, 2023, 100160, ISSN 2666-920X, https://doi.org/10.1016/j.caeai.2023.100160.

Concerns from instructors:

- How about the performance of these advanced AI tools?
- Using these advanced AI tools, can students practice higher-order thinking (e.g., independent thinking)?
- How can we better design tasks and guide students to use these advanced AI tools?

Challenges and Our Approaches

- Lack of dataset -> Integration of ChatGPT
- Lack of analysis from cognitive levels -> Creation with pedagogical insights
- Lack of ability to track the progression of the various LLMs' responses and observe how students adjust their prompts in response -> Visual analytics system (StuGPTViz)

Integration of ChatGPT in Education

- We integrated ChatGPT into the curriculum of a postgraduate data visualization course for computer science majors in the first semester of 2024.
- Each in-class exercise session, we conducted the experiment during the last 40 minutes of the lecture, included a 10-minute self-learning segment with ChatGPT, a 25-minute task completion segment, and a 5-minute conversation log upload phase.
- 744 unique conversations with 2507 turns after filtering out the empty conversations and those unrelated to the learning tasks

Dataset Creation with Pedagogical Insights

Task Type & Count	Task Brief	Cognitive Level
Concept Remember (2)	Multiple Choices questions for basic concept remembering	**Remember (L1)**
Concept Understanding (3)	Multiple Choices questions for deeper concept understanding	**Understand (L2)**
Concept Application (3)	Short questions for concept application	**Apply (L3)**
Visualization Analysis (4)	Open-ended analysis questions (e.g., encoding usage, color scheme)	**Analyze (L4)**
Visualization Evaluation (5)	Evaluate the given visualization design	**Evaluate (L5)**
Visualization Design (4)	Design visualization with the given data	**Create (L6)**
Self Learning (6)	Self exploration of key concepts	**Others**

StuGPTViz: Visual Analytics System

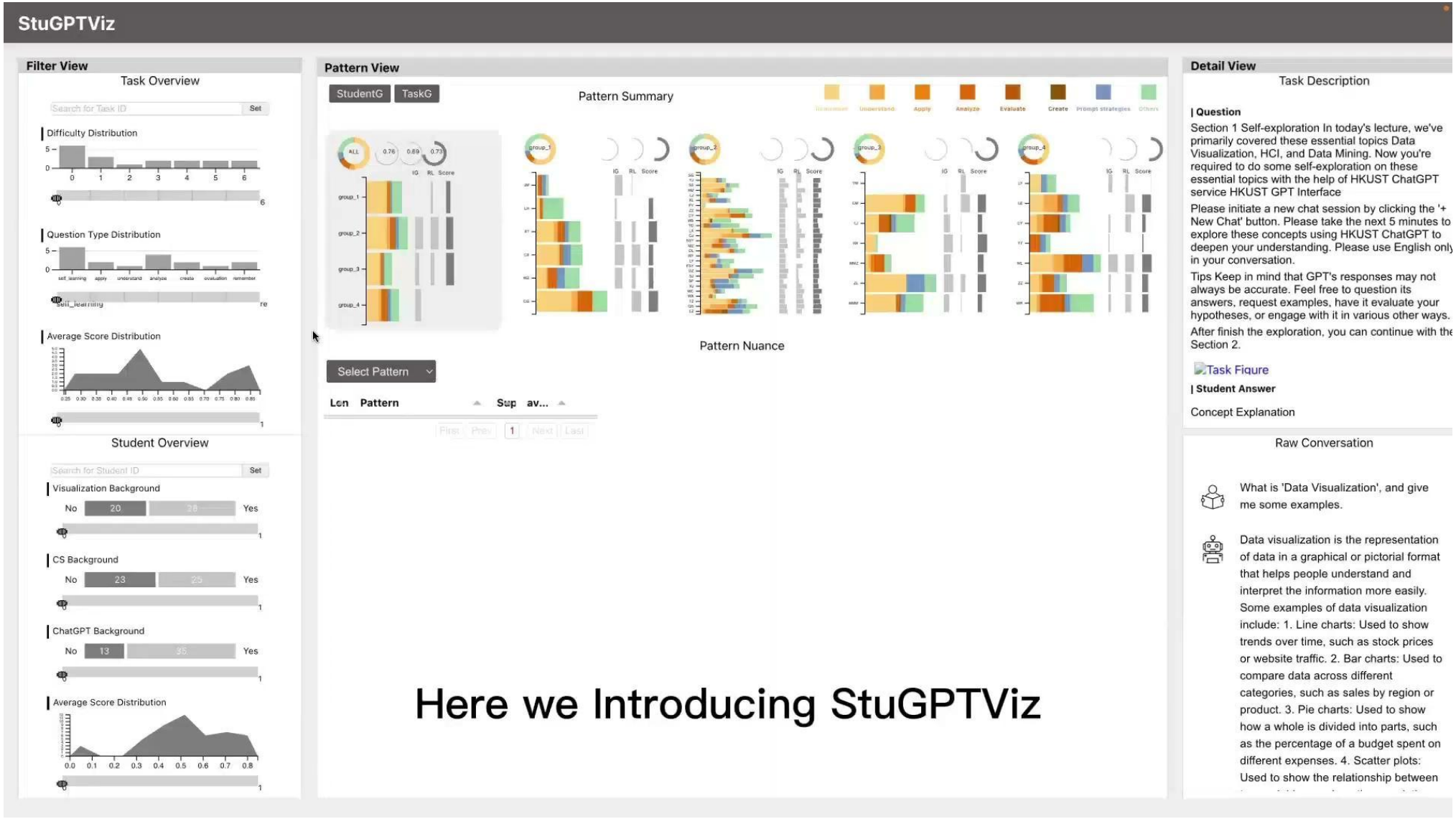
R1: Overview of students and tasks data

R2: Summarizing macro-level conversation characteristics

R3: Identifying micro-level interaction patterns

R4: Tracing interaction pattern evolution

R5: Evaluating interaction pattern performance


StuGPTViz
Filter View
Task Overview
Search for Task ID
Set
Difficulty Distribution
Question Type Distribution
Average Score Distribution
Student Overview
Search for Student ID
Set
Visualization Background
No
Yes
CS Background
No
Yes
ChatGPT Background
No
Yes
Average Score Distribution
Pattern View
StudentG
TaskG
Pattern Summary
Pattern Nuance
Select Pattern
Len
Pattern
Sup
av...
Here we Introducing StuGPTViz
Detail View
Task Description
| Question
Section 1 Self-exploration In today's lecture, we've primarily covered these essential topics Data Visualization, HCI, and Data Mining. Now you're required to do some self-exploration on these essential topics with the help of HKUST ChatGPT service HKUST GPT Interface
Please initiate a new chat session by clicking the '+ New Chat' button. Please take the next 5 minutes to explore these concepts using HKUST ChatGPT to deepen your understanding. Please use English only in your conversation.
Tips Keep in mind that GPT's responses may not always be accurate. Feel free to question its answers, request examples, have it evaluate your hypotheses, or engage with it in various other ways.
After finish the exploration, you can continue with the Section 2.
Task Figure
| Student Answer
Concept Explanation
Raw Conversation
What is 'Data Visualization', and give me some examples.
Data visualization is the representation of data in a graphical or pictorial format that helps people understand and interpret the information more easily. Some examples of data visualization include: 1. Line charts: Used to show trends over time, such as stock prices or website traffic. 2. Bar charts: Used to compare data across different categories, such as sales by region or product. 3. Pie charts: Used to show how a whole is divided into parts, such as the percentage of a budget spent on different expenses. 4. Scatter plots: Used to show the relationship between

Evaluation and Result

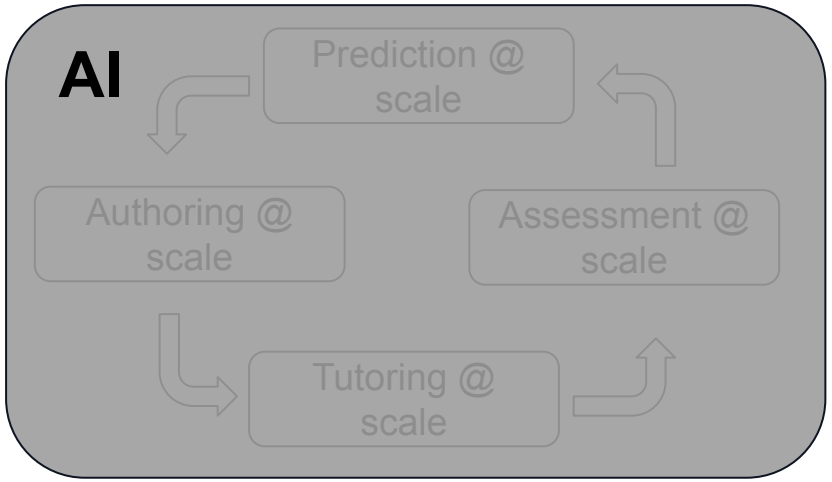
- Students' learning perspective: Questionnaire Feedback
  More than 90% students enjoy using ChatGPT in their learning process
- ChatGPT performance:
  Strong positive correlation between the IG (information gain) metric and experts' judgment of ChatGPT's response quality
- Expert interviews:
  "The ability to discern students' overall cognitive level at a glance is highly appreciated."

  "The workflow's logical progression and the interconnection of each view were particularly impressive, enabling a diverse analytical focus through a unified procedure."

Human

Content inaccuracy
Improper use of AI

Analyzing @ scale

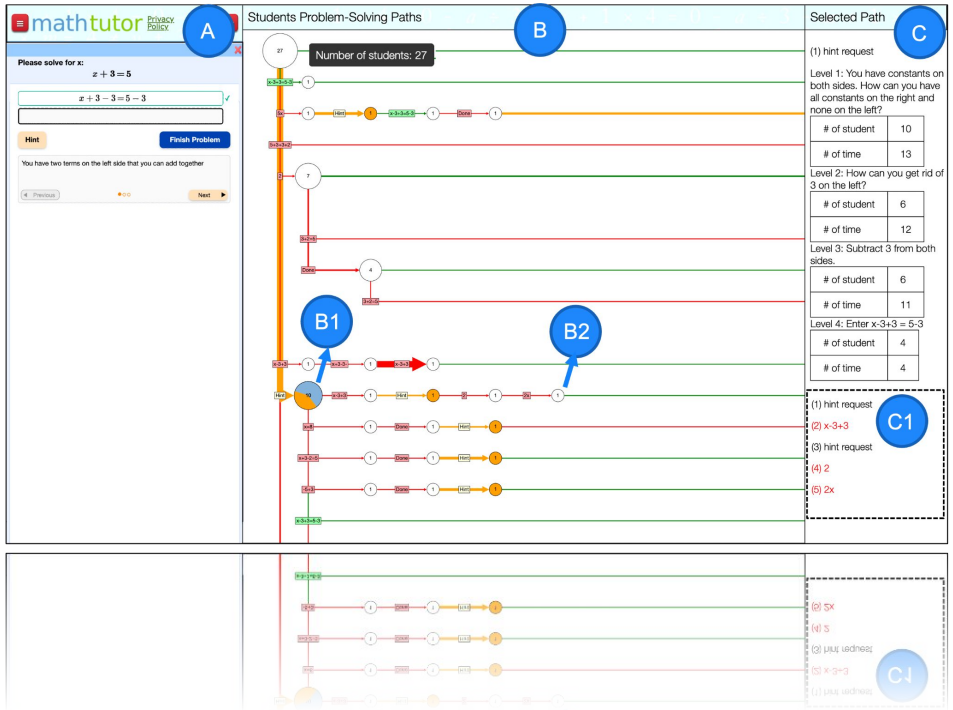


Lack of pedagogical guidance
Not personalized

Explaining @ scale

Lack of evaluation
High risk

Testing @ scale



Involving Teachers in the Data-driven Improvement of Intelligent Tutors: A Prototyping Study

Meng Xia, Xinyi Zhao, Dong Sun, Yun Huang, Jonathan Sewall, Vincent Aleven

AIED 2023

Carnegie Mellon University

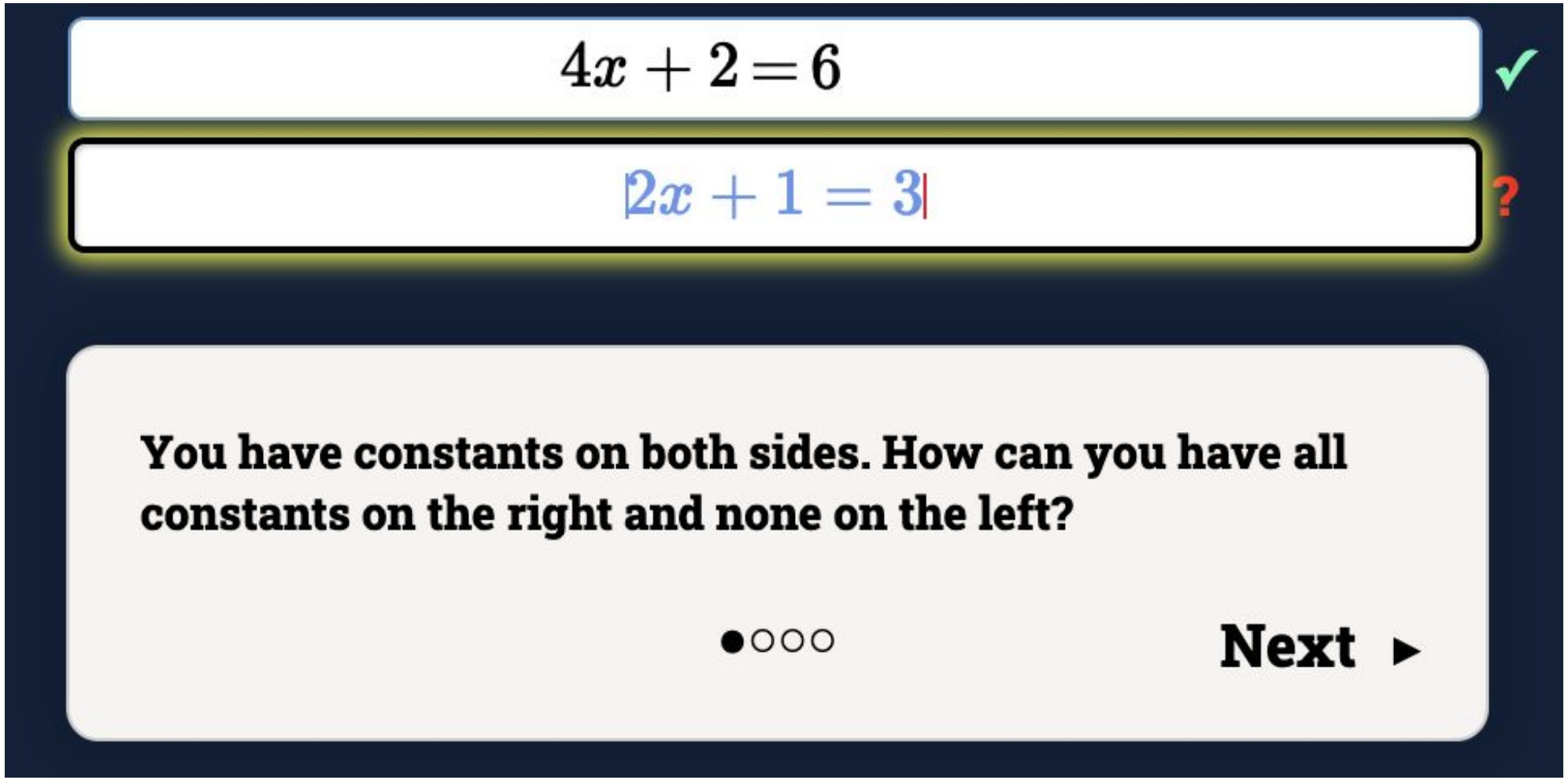
FUDAN UNIVERSITY

NIO

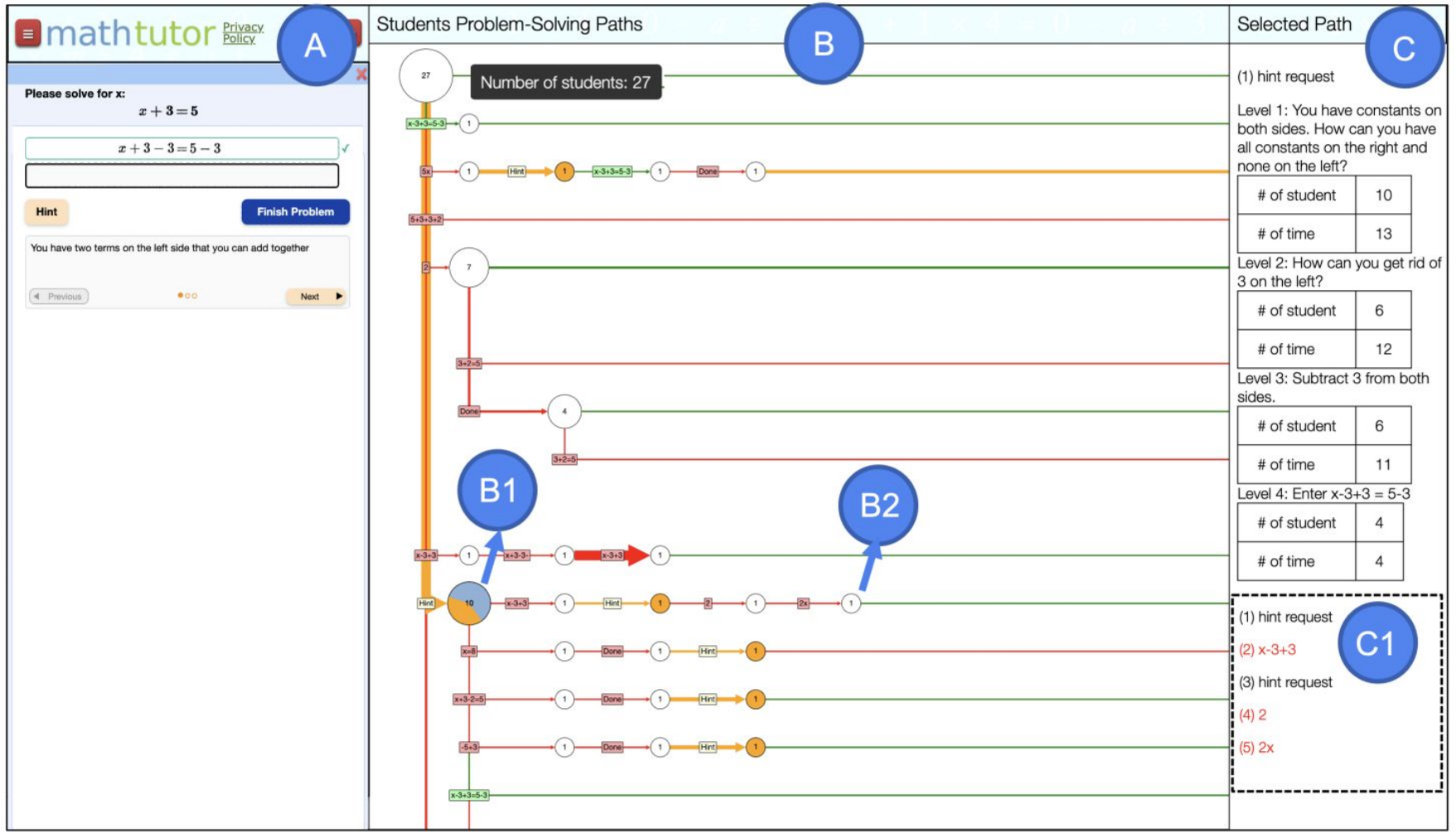
AI's role: tutoring @ scale

Vis' role: explaining @ scale

Initial design of intelligent tutoring systems often not optimal!



Research Prototype: SolutionVis

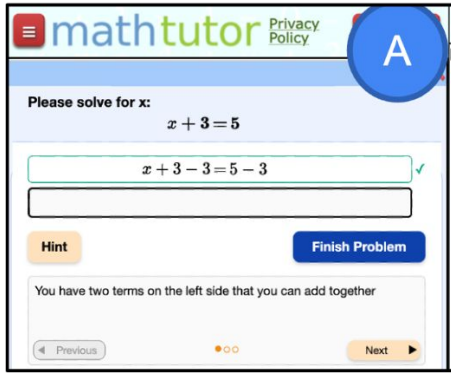


User Study

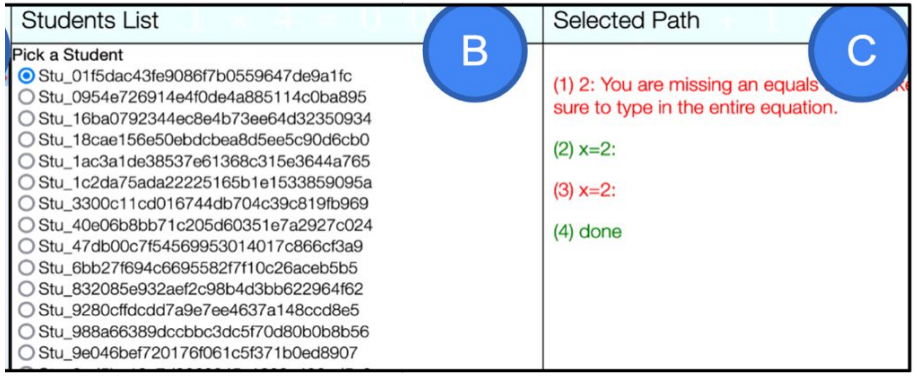
Participants: Eight middle school math teachers

Task: Explore different interfaces and give suggestions on how to improve the intelligent tutor.

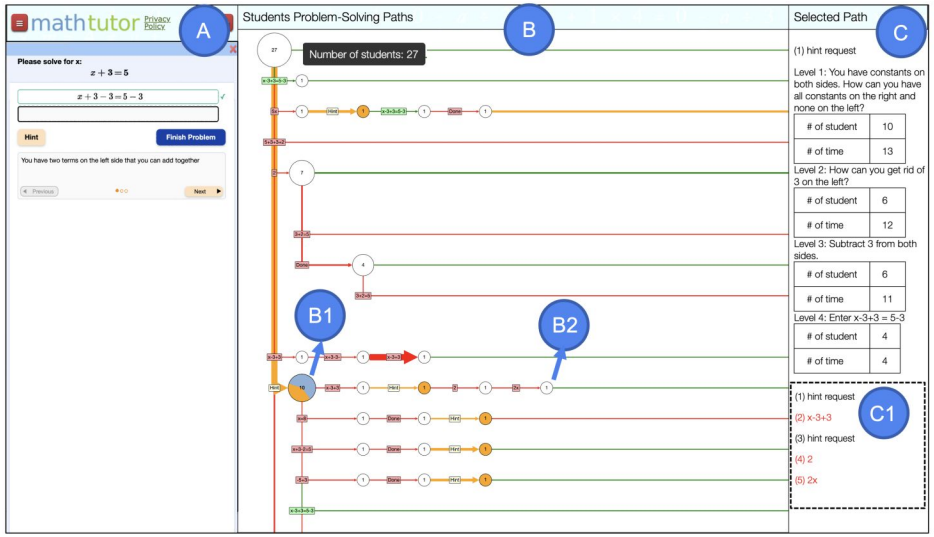
Conditions:



"no data" Baseline



List Interface



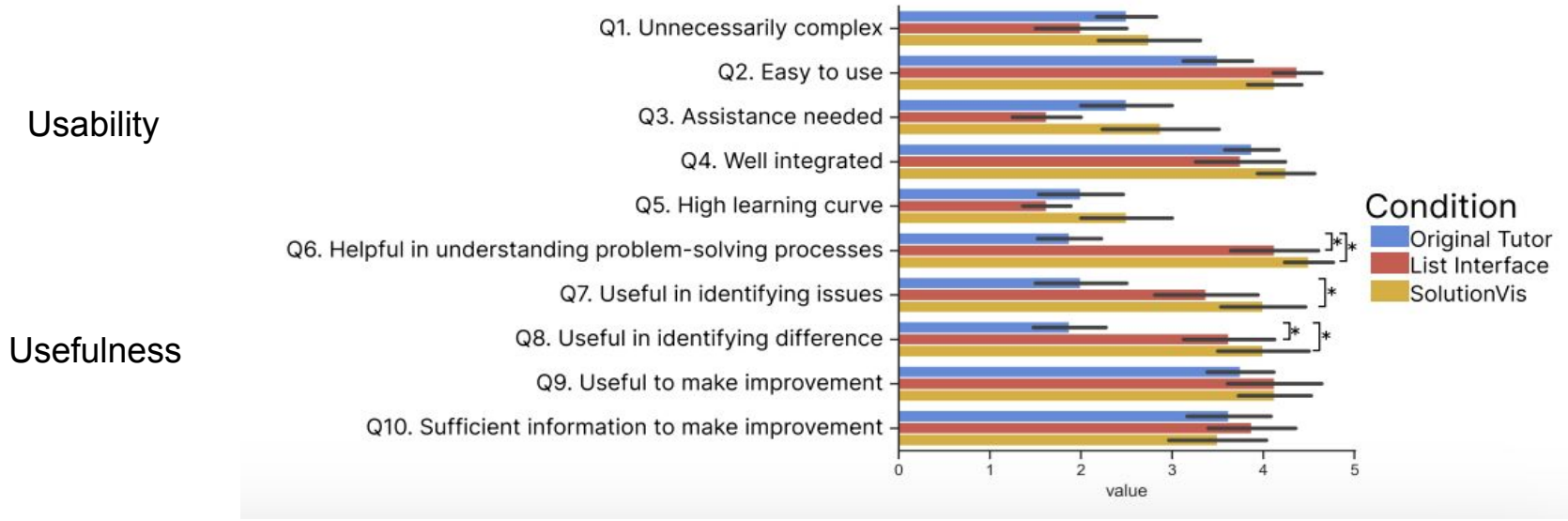
SolutionVis

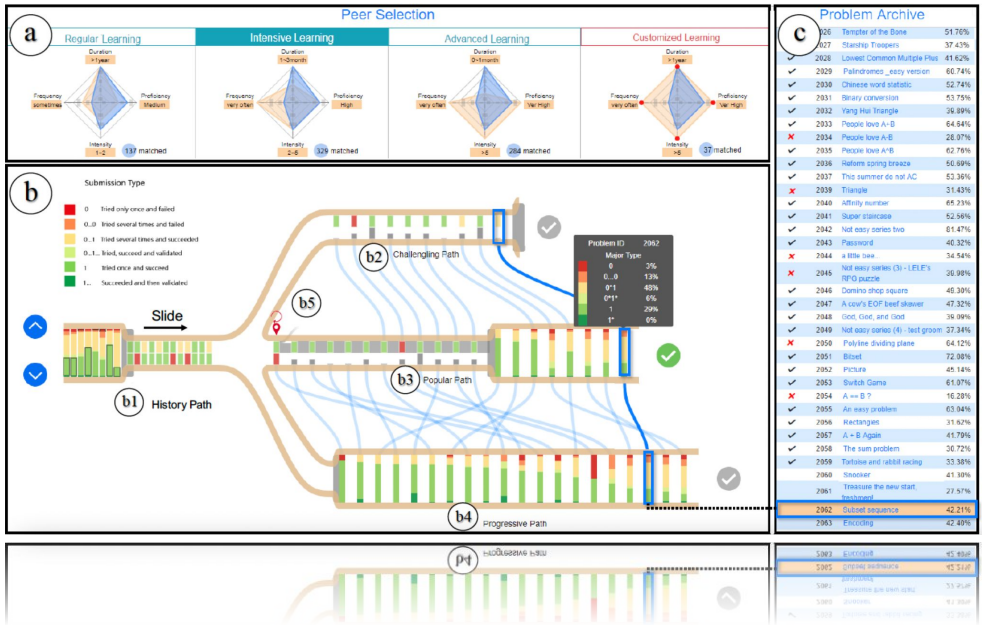
Feedback

		Original Tutor	List Interface	SolutionVis
Interface/logic Design		Need to provide instruction that using "Enter" instead of "Finish problem" to go to next step. (5) The correct step with a different order of the items in the equation is not accepted by the tutor. (5) Cursor should go to the next line automatically.(1) Need to say explicitly to input the final answer or the intermediate step.(2)		
Hints Design	Hint Adaptivity	Address what the students did in the hint. (3)	If the answer is super close to the final answer (e.g., missing a negative sign), provide more concrete feedback (e.g., check your sign) (1)	Provide hints based on how many times the student asks for it. If the student asks a second time, showing a different hint. (1)
	Hint Clarity/ Correctness	The hint "You can get the variable by itself by dividing both sides by the coefficient." is not correct for "-2x-4=2". Should "add 4 on both sides first".(3) Rewrite the hint "Your input in not valid algebra." as "You need to have 'x ='"(2) The hint "You have two terms on the left side that you can add together" is not well designed. Explain and give examples about "term".(4)	Provide the number line in the hint for students to understand positive and negative numbers for the steps where they need to move items from one side to the other side. (1)	Check the first hint of each step and make sure it is clear and easy to understand. (1)
	Hint Visibility	Don't show the bottom hint.(1) Let the hints pop up automatically.(1) Show part of the hints to let students think more at each step; teaching them but not correcting them.(1)	Don't encourage students to ask for hints at beginning, but ask "what would be your first step?" (1)	
	Hint Composition	—	—	Ask a question about the knowledge in the hint to let them think. (1) Show some exmaples in the hint. (1)
Feedback Design: Gaming the System/ Protracted Struggle		—	When seeing students gaming the system or inputting random things, provide feedback like "show me your efforts", "show your work"(2) When seeing student submitting the same thing multiple times, let the tutor give the answer and move on.(1)	When seeing students gaming the system or inputting random things, provide feedback like "show me your efforts", "show your work" (4) When seeing student submitting the same thing multiple times, provide feedback like “stop clicking this, please try again” (2) Show cartons or funny animals to encourage them.(1)

Results

1. Data about student learning was helpful for teachers to generate useful redesign ideas.
2. The aggregated data in a graph showing in SolutionVis helps teachers find the tutor's problems efficiently





PeerLens: Peer-inspired Interactive Learning Path Planning in Online Question Pool

Meng Xia, Mingfei Sun, Huan Wei, Qing Chen, Yong Wang, Lei Shi, Huamin Qu, Xiaojuan Ma

CHI 2019

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

AI's role: prediction @ scale

Vis' role: explaining @ scale

Motivation

Pro. ID	
1000	A + B Problem
1001	Sum Problem
1002	A + B Problem II
1003	Max Sum
1004	Let the Balloon Rise
1005	Number Sequence
1006	Tick and Tick
1007	Quoit Design
1008	Elevator
1009	FatMouse' Trade
1010	Tempter of the Bone
1011	Starship Troopers

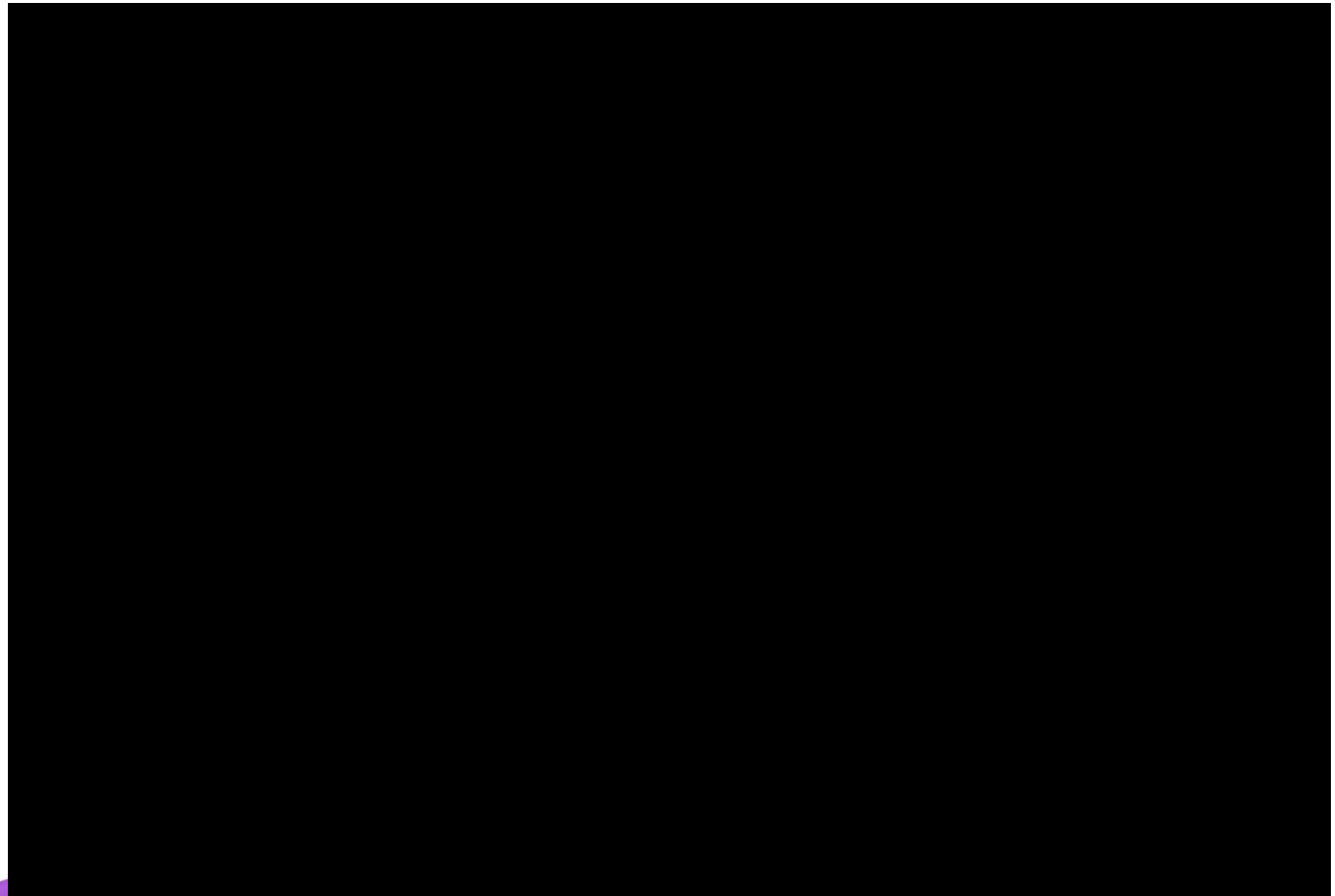
Questions Pools:

- No pre-determined syllabus
- A lengthy list indexed by their problem IDs
- Hidden intents

Learners:

- Different learning scenarios
- One learner's learning scenario may be changing

What to do next? What sequence to follow?



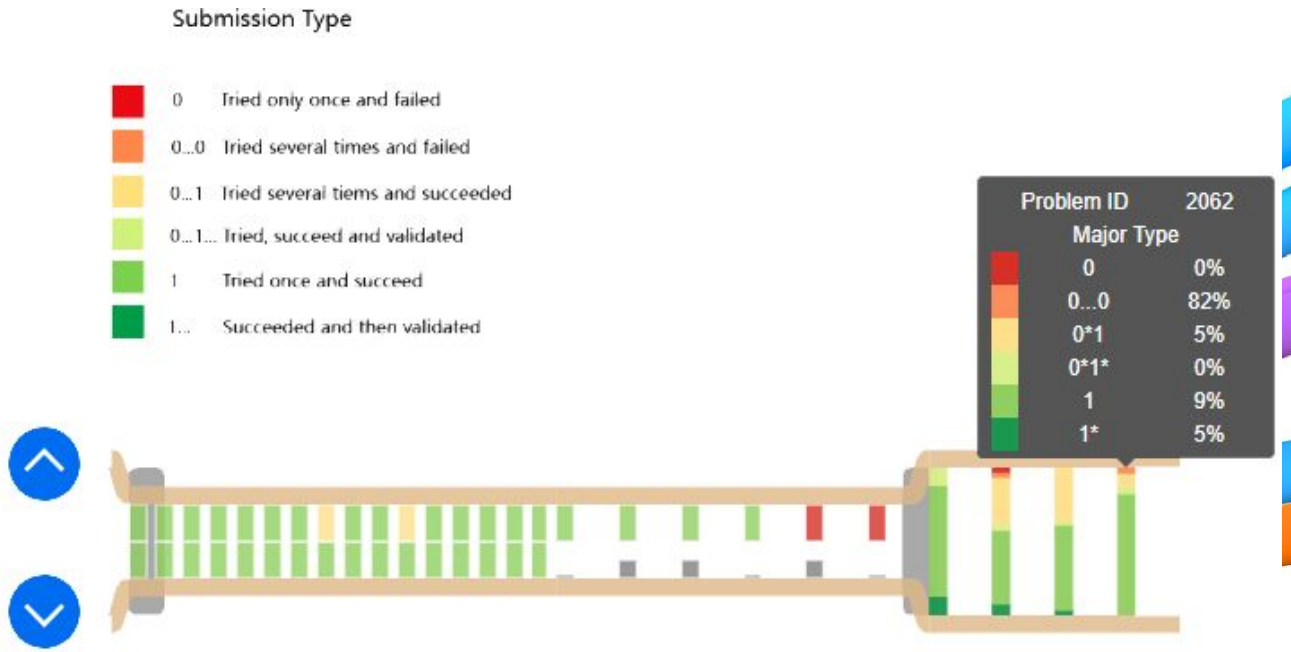
Evaluation: Controlled User Study

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15···33 34 35 36 37 38
39 40 41 42 43 44 45 46 47 48 49 50

Search: In Title Go

Pro. ID	Problem Title	Ratio(Accepted/Submissions)
1000	A + B Problem	30.56%(240770/787844)
1001	Sum Problem	25.38%(143110/563922)
1002	A + B Problem II	19.47%(84152/432201)
1003	Max Sum	23.76%(70413/296345)
1004	Let the Balloon Rise	39.72%(59043/148661)
1005	Number Sequence	25.25%(51499/203970)
1006	Tick and Tick	26.73%(6080/22750)
1007	Quoit Design	26.52%(17197/64856)
1008	Elevator	54.79%(46878/85565)
1009	FatMouse' Trade	34.85%(33070/94883)
1010	Tempter of the Bone	26.68%(39786/149139)

Baseline system (List View)



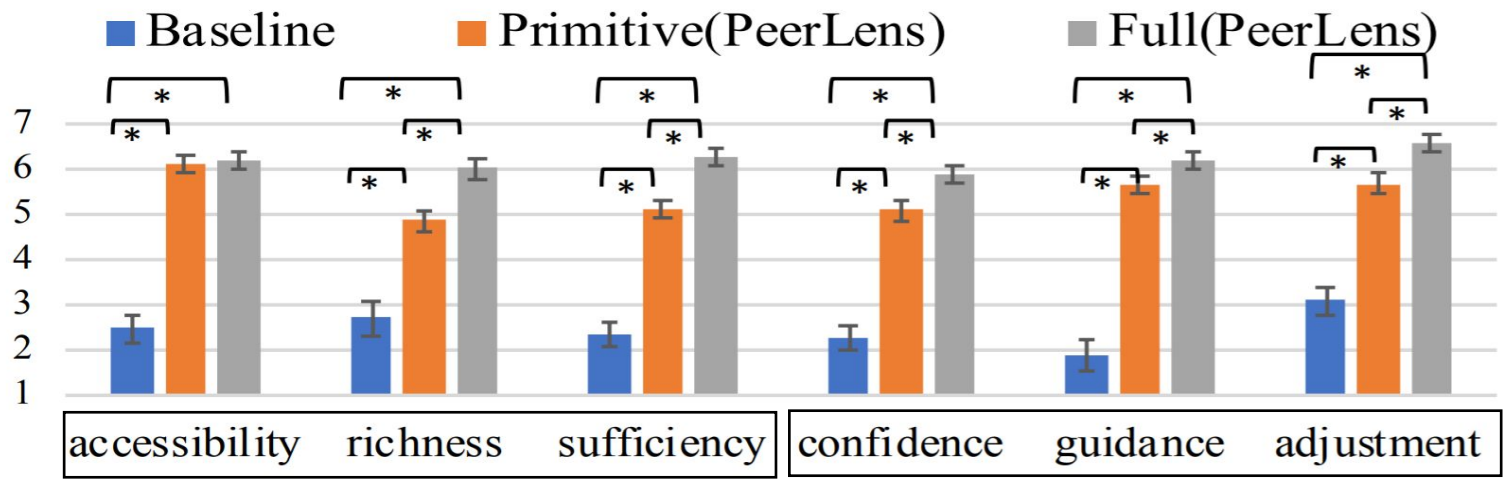
Primitive PeerLens (Only provide one path)

18 CS students:

- determine the starting question under a specific learning scenario
- find the next question to solve given an existing historical learning path

Results

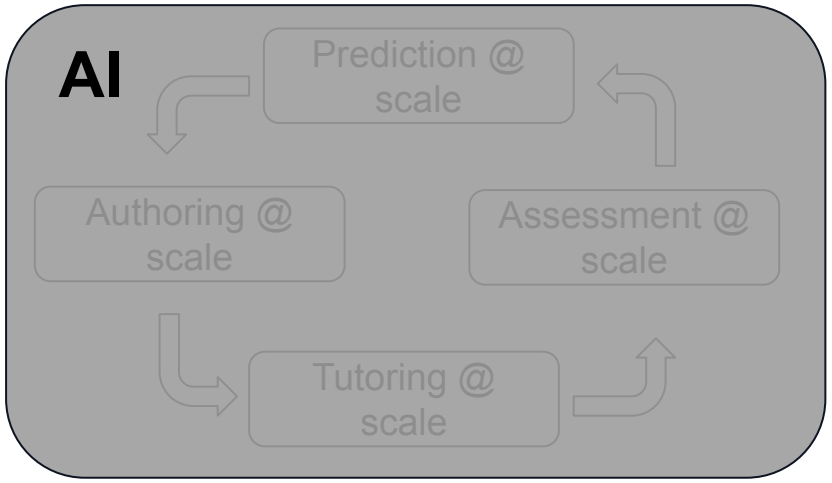
1. Using peer data is useful.
2. Visualizing more suggestions is useful.
3. Visualizing more suggestions using the proposed visualizations did not increase the complexity.



Human

Content inaccuracy
Improper use of AI

Analyzing @ scale

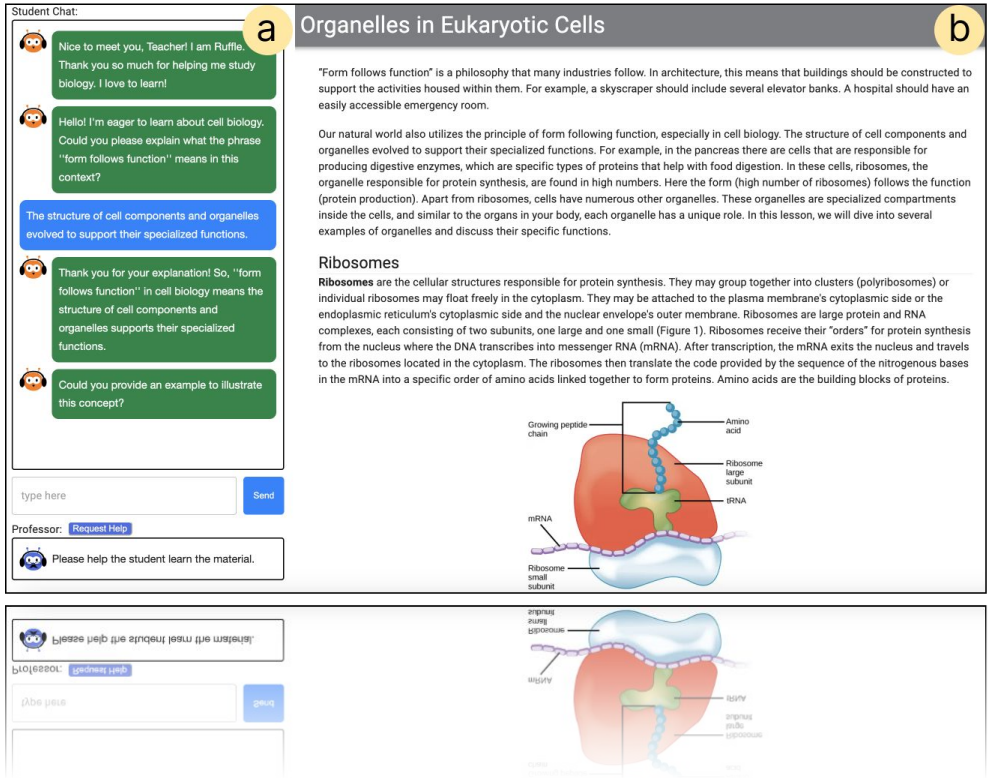


Lack of pedagogical guidance
Not personalized

Explaining @ scale

Lack of evaluation
High risk

Testing @ scale



Ruffle&Riley: Insights From Designing and Evaluating a LLM-Based Conversational Tutoring System

Robin Schmucker, Meng Xia, Amos Azaria, Tom Mitchell

AIED 2024

Carnegie Mellon University
ML Machine Learning Department
ATM
Ariel University

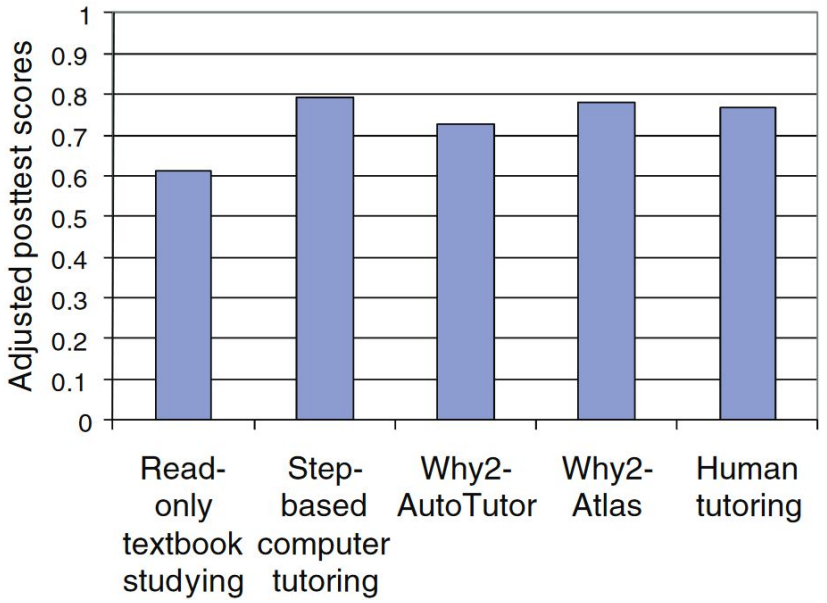
AI's role: authoring/tutoring/grading @ scale

Vis' role: testing @ scale

Intelligent Tutoring Systems (ITSs)

Benefits of ITSs:

- Can be as effective as human tutoring
- Can be scaled to millions of learners
- Provide a formal framework for thinking about tutoring processes



Learning outcomes of different activities [1]

[1] VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational psychologist*, *46*(4), 197-221.

Intelligent Tutoring Systems (ITSs)

Limitations of ITSs:

- High cost of content authoring
- Limited language understanding
- Limited conversational facilities
- Limited question answering facilities

LLMs help us address these challenges


Chapter 10
AutoTutor
Arthur C. Graesser
Zhiqiang Cai
USA
USA
USA
dialogues are organized around difficult questions and problems that require reasoning and explanations in the answers. The major components of AutoTutor include an animated conversational agent, dialogue management, speech act classification, a curriculum script, semantic evaluation of student contributions, and electronic documents (e.g., textbook and glossary). This chapter describes the computational components of AutoTutor, the similarity of these components to human tutors, and some challenges in handling smooth dialogue. We describe some ways that AutoTutor has been evaluated with respect to learning gains, conversation quality, and learner impressions. AutoTutor is sufficiently modular that

AutoTutor Recap [1]

[1] Graesser et al. (2012). AutoTutor. *Applied Natural Language Processing: Identification, Investigation and Resolution*, 169-187.

How can LLMs benefit ITSs?

LLMs can provide components to ITS designers:

- By generating lesson texts
- By generating question and hints
- By adding automatic grading capabilities
- By adding question answering capabilities
- …

What would it take to generate an **entire** ITS?

Ruffle&Riley: User Interface

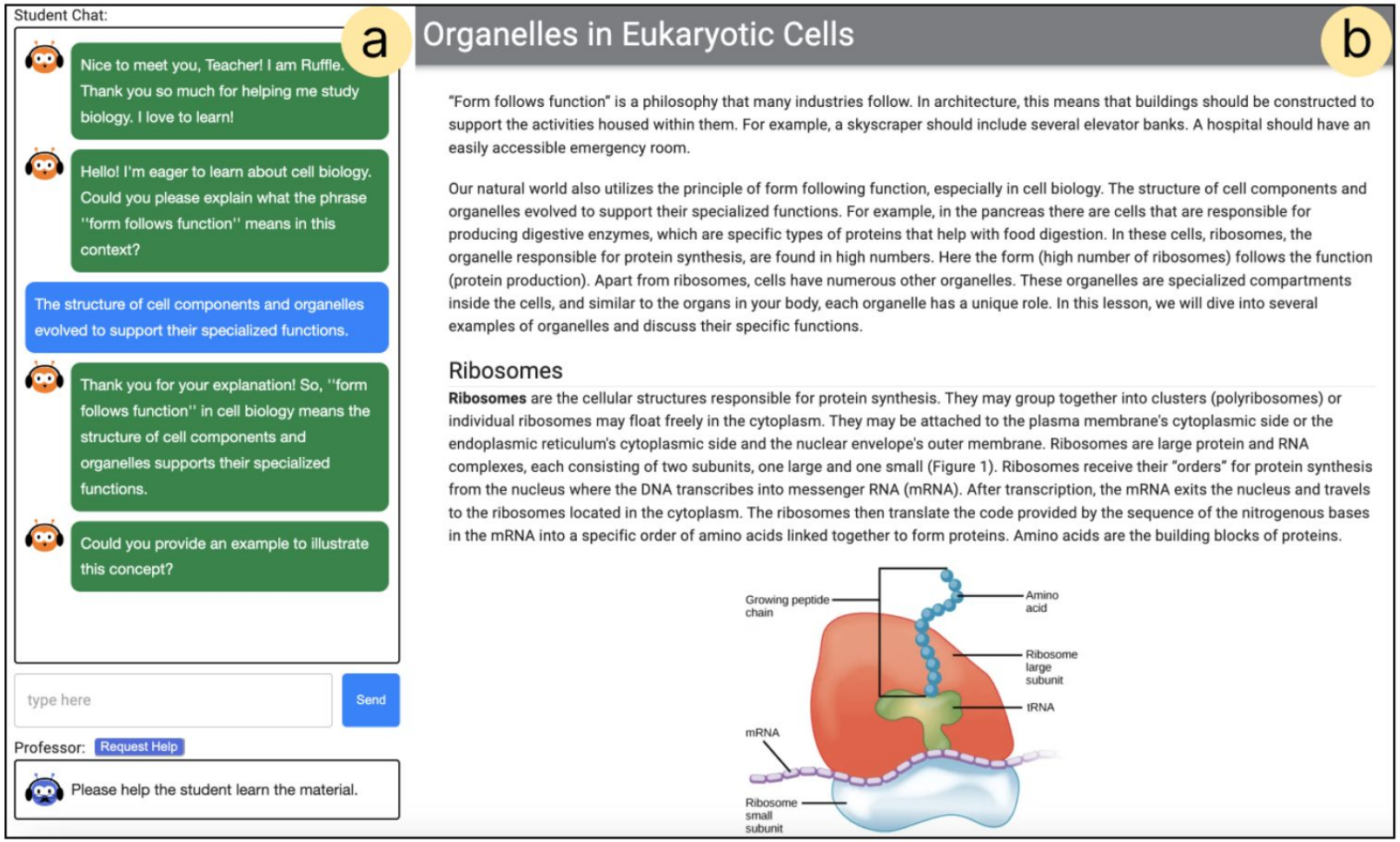
Free-form dialog

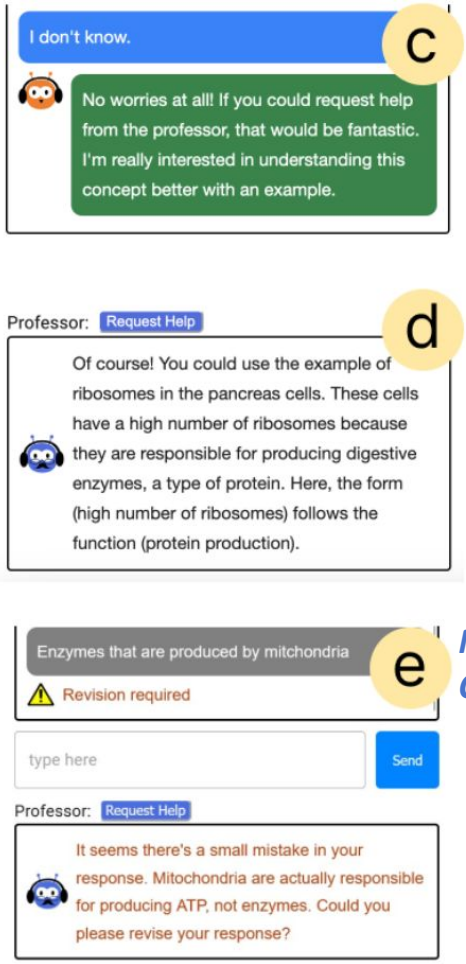
Pre-existing textbook content

Intelligent Feedback

On demand assistance

Misconcept. Correction





User Interface

Evaluation Results

Findings of user study

Learning Experience Survey

Conditions	Learning Experience (1-strongly disagree, 7-strongly agree)						
	Engagement	Understanding	Remembering	Interruption	Coherence	Support	Enjoyment
Reading	4.33 ± 0.52	-	-	-	-	-	-
Teacher Q/A	5.0 ± 0.53	4.43 ± 0.65 *	4.43 ± 0.65 *	2.71 ± 0.64	5.43 ± 0.53	4.57 ± 0.57 *	3.71 ± 0.52 *
LLM Q/A	4.8 ± 0.47	4.4 ± 0.4 *	4.33 ± 0.42*	2.67 ± 0.45	4.8 ± 0.43*	4.0 ± 0.44 *	4.0 ± 0.44 *
Ruffle & Riley	**5.81 ± 0.3**	**5.81 ± 0.24**	**5.76 ± 0.22**	**2.19 ± 0.34**	**6.1 ± 0.21**	**5.9 ± 0.26**	**5.62 ± 0.31**

Post-Test Performance

Conditions	Previous Knowledge	Learning Performance
		Post-test Scores (i.e.,Multiple-Choice Questions)
Reading	2.53 ± 0.41	5.07 ± 0.33
Teacher Q/A	**3.0 ± 0.58**	4.14 ± 0.83
LLM Q/A	2.2 ± 0.3	4.67 ± 0.35
Ruffle & Riley	2.67 ± 0.43	**5.19 ± 0.25**

Symbol "" marks $p < 0.05$*

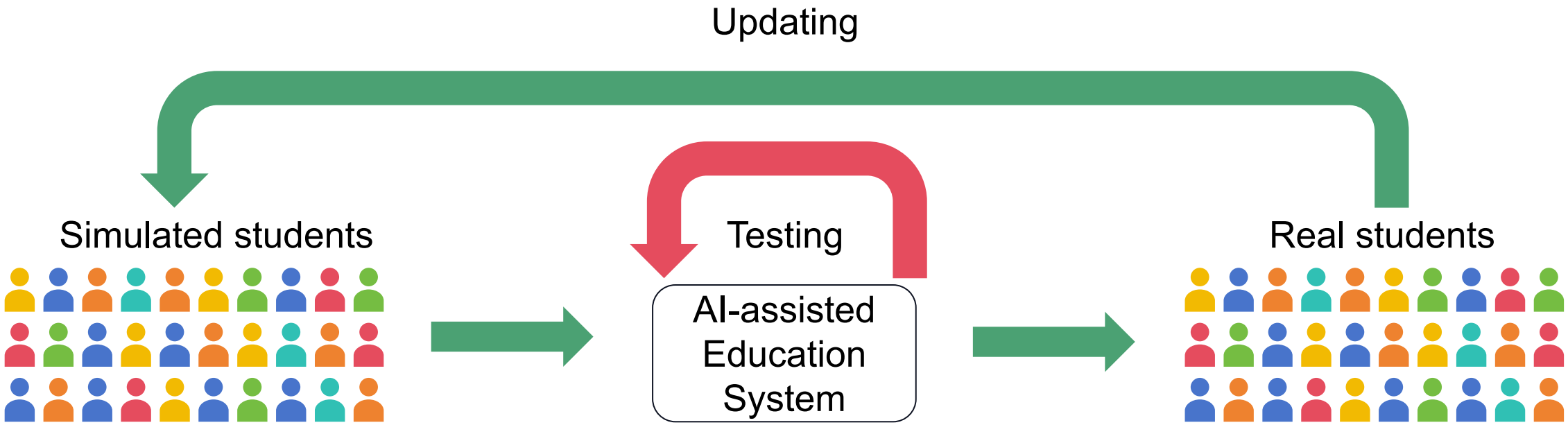
Interaction Analysis

How successful is Ruffle&Riley at orchestrating conversational tutoring?

- Found no “hallucination” in GPT-4’s outputs
- Sometimes asks for previously covered information
- Lenient towards incomplete explanations

Next step: Simulation & Digital Twin

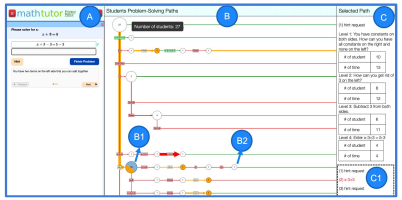
We are now working how to simulate different learning scenarios and provide educators the feedback at scale using visual analytics so that they can confidently deploy the system in reality.



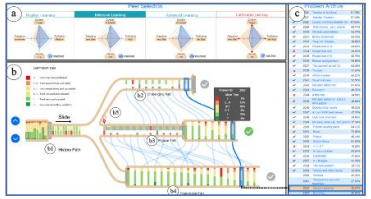
Analyzing @ scale

StuGPTVis (TVCG 2024)

Explaining @ scale

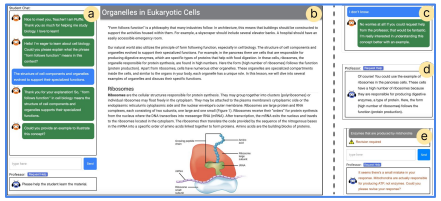


SolutionVis (AIED 2023)



Peerlens (CHI 2019)

Testing @ scale



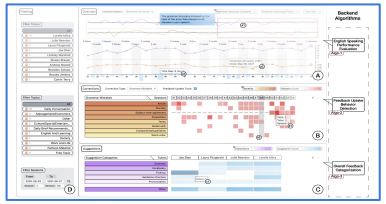
Ruffle&Riley (AIED 2024)

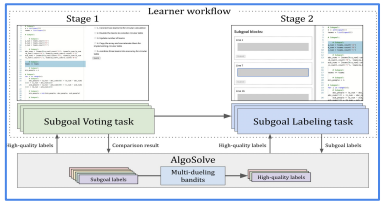
AI's role: tutoring @ scale tutoring/prediction @ scale authoring/tutoring @ scale

Vis' role: analyzing @ scale explaining @ scale testing @ scale

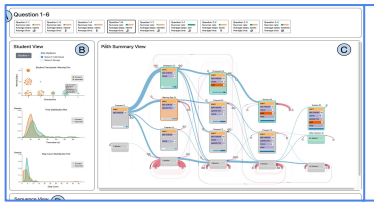
Analyzing @ scale

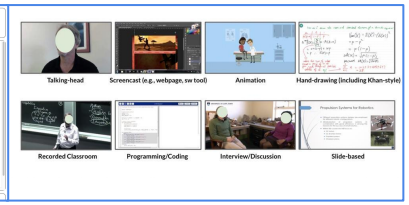

StuGPTVis (TVCG 2024)


RLens (L@S 2022)

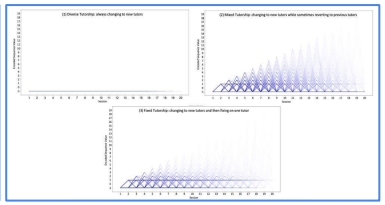


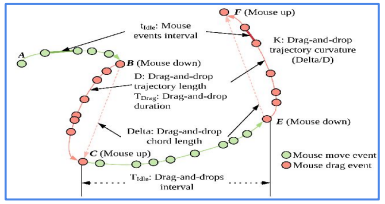
AlgoSolve (CHI 2022)


BlockLens (L@S 2022)

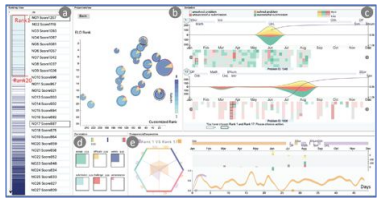


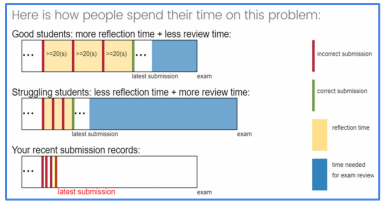
Mobile MOOCs (CHI 2022, **Best Paper Award**)


Distributed Tutorship (LAK 2022)

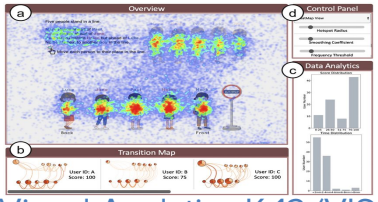


Predication (LAK 2020)


SeqDynamics (EuroVIS 2020)

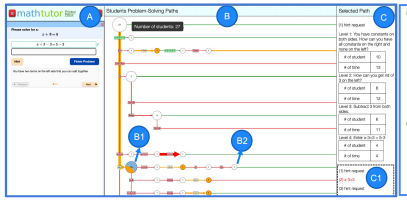


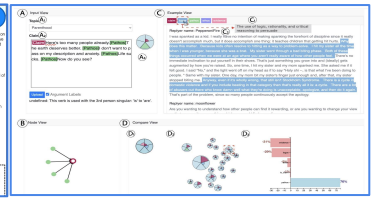
"Gaming the system" (L@S 2020)



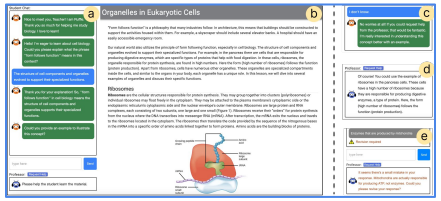
Visual Analytics K-12 (VIS 2019, **Best Poster Award**)

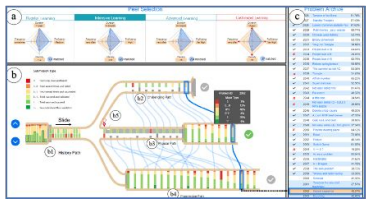
Explaining @ scale


SolutionVis (AIED 2023)

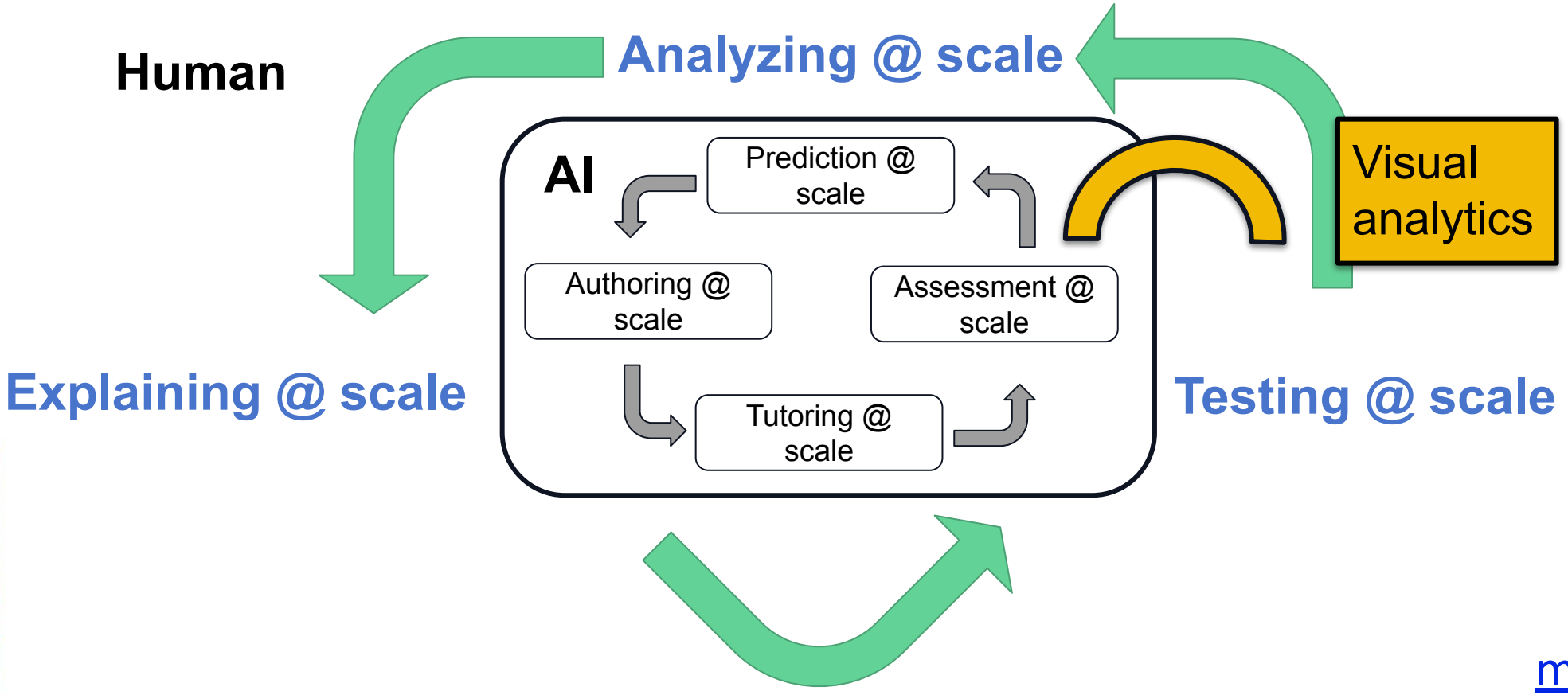

Persua (CSCW 2022)

Testing @ scale


Ruffle&Riley (AIED 2024)


Peerlens (CHI 2019)


QLens (TVCG 2021)



Meng Xia
mengxia@tamu.edu
www.mengxia.org

Use **visual analytics**, **AI**, and other **human-AI interaction techniques** and research metaphors to **upskill** educators and learners to better utilize data and AI for **Personalization@Scale**!

DLab Dream Lab
TEXAS A&M UNIVERSITY

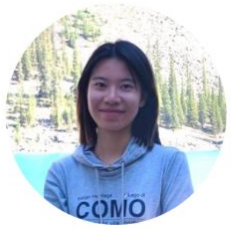
DL Dream Lab

News People Publications Contact

Graduate Students

Jiwon Chun

Master Student (2024 Spring)



Yi Wen

PhD Student (2024 Fall)

Fatemeh Mirhosseini

PhD Student (2024 Fall)

Hangxiao Zhu

PhD Student (2024 Fall)

Visiting Students

Gefei Zhang

Shenming Ji

Sitong Pan

https://www.xiameng.org/DreamLab/