Human-AI Interaction for Personalized Online Learning

Meng Xia

April 10, 2024
Online Learning has become widespread

A Coding Exercises Platform

A MOOCs platform (massive open online courses)

A videotelephony software

An online discussion forum

Virtual teacher (Chatbot)
Learning at scale, flexible locations and times

Scale

Space

Time

Machine Learning Specialization

#BreakIntoAI with Machine Learning Specialization. Master fundamental AI concepts and develop practical machine learning skills in the beginner-friendly, 3-course program by AI visionary Andrew Ng

🌟🌟🌟🌟 4.9 7,331 ratings

Andrew Ng +3 more instructors  TOP INSTRUCTORS

120,420 already enrolled
Learning at scale, flexible locations and times

- Scale
- Space
- Time

Make learning tailored to individual students a grand challenge!
Why personalized learning?

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)

Learning at scale, flexible locations and times

Scale

Space

Time

How could we achieve personalized online learning?
Learning Data is a Communication Channel

e.g. Mouse movement data
Online Learning Platforms and Data

Problem-solving data

Click-stream data

Video and audio communication data

Online discussion data
Help learners and educators to find insights from learning data and use the insights to make decisions for achieving personalized online learning.
Related Work on Learning Analytics

● Focus more on cognitive skills instead of non-cognitive variables (e.g., motivations, self-regulation skills)
  Learning factor analysis (Cen et al., 2006) Performance factor analysis (Pavlik et al., 2009)

● Algorithms’ outputs are not easy to interpret
  Deep knowledge tracing (Piech et al., 2015) Explanatory models for educational data (Liu et al., 2017)
My Approach

Design and build human-centered data-driven systems for achieving personalized online learning.
Why *human-centered* data-driven systems?

Learners’ needs  Educators’ domain knowledge

Data  AI Algorithm  Human  Decision Making

data-driven interfaces (e.g., data visualization)
Challenges

- **Learners have different aspects** that change over time

- **Learners and educators are not data analysts**

- Different learning scenarios:
  - Learning data is **voluminous and heterogeneous**
  - Learners and educators have **different tasks**
Learners (learning loop)

Planning

Performing

Data

Reflecting

Understanding and Analyzing

Redesigning

Educators (design loop)

PeerLens (CHI 2019)
Persua (CSCW 2022)
RLens (L@S 2022)
StuGPTVis (TVCG 2024)
QLens (TVCG 2021)
SolutionVis (AIED 2023)

Predication (LAK 2020)
AlgoSolve (CHI 2022)
"Gaming the system" (L@S 2020)
Mobile MOOCs (CHI 2022, Best Paper Award)
Visual Analytics K-12 (VIS 2019, Best Poster Award)

Ruffle&Riley (AIED 2024)
Distributed Tutorship (LAK 2022)
SeqDynamics (EuroVIS 2020)
BlockLens (L@S 2022)
Learners (learning loop)

Planning
- Data

Performing
- Peerlens (CHI 2019)
- Persua (CSCW 2022)
- AlgoSolve (CHI 2022)
- Ruffle&Riley (AIED 2024)

Reflecting
- Predication (LAK 2020)
- "Gaming the system" (L@S 2020)
- Distributed Tutorship (LAK 2022)

Understanding and Analyzing

Redesigning
- QLens (TVCG 2021)
- BlockLens (L@S 2022)

Educators (design loop)

- RLens (L@S 2022)
- AlgoSolve (CHI 2022)
- Persua (CSCW 2022)
- "Gaming the system" (L@S 2020)
- Predication (LAK 2020)
- Distributed Tutorship (LAK 2022)

- Peerlens (CHI 2019)
- Persua (CSCW 2022)
- AlgoSolve (CHI 2022)
- Ruffle&Riley (AIED 2024)

Mobile MOOCs (CHI 2022, Best Paper Award)
Visual Analytics K-12 (VIS 2019, Best Poster Award)
StuGPTVis (TVCG 2024)
SolutionVis (AIED 2023)
SeqDynamics (EuroVIS 2020)
Ruffle&Riley (AIED 2024)
Learners (learning loop)

Planning

Performing

Reflecting

Understanding and Analyzing

Redesigning

Data

Planning

Performing

Reflecting

Understanding and Analyzing

Redesigning

Peerlens (CHI 2019)
Persua (CSCW 2022)
RLens (L@S 2022)

Predication (LAK 2020)
AlgoSolve (CHI 2022)
“Gaming the system” (L@S 2020)

Ruffle&Riley (AIED 2024)
Distributed Tutorship (LAK 2022)

StuGPTVis (TVCG 2024)
QLens (TVCG 2021)
SolutionVis (AIED 2023)

Mobile MOOCs (CHI 2022, Best Paper Award)
Visual Analytics K-12 (VIS 2019, Best Poster Award)

SeqDynamics (EuroVIS 2020)
BlockLens (L@S 2022)

How data can be used for learners to plan, perform, and reflect on their learning?

How data can be used for educators to improve the design of learning materials?

How data can be used for learners to plan, perform, and reflect on their learning?

How data can be used for educators to improve the design of learning materials?

Peerlens (CHI 2019)  
Persua (CSCW 2022)

QLens (TVCG 2021)  
StuGPTVis (TVCG 2024)
How data can be used for learners to **plan**, perform, and reflect on their learning?
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
What is an online question pool?

- A collection of questions for learners to practice their knowledge online
Motivation

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?
A user-centered design process

- Participants: question pool designers and question pool users
- Requirements gathering iteratively for three months
  - R1: Find peers for a target learning scenario.
  - R2: Compare with peers’ performance to understand the gap.
  - R3: Offer flexible learning path suggestions with explanations.
  - R4: Provide convenient interaction and intuitive visual designs.
PeerLens System overview

1. Dataset
   - Crawling

2. Preprocessing
   - Filtering
   - Statistics Calculation
   - Sequence Modeling

3. Path Engine
   - Peers Selection
   - Path Planning
   - Prediction and Hints

4. Visualization
   - Basic Info
     - User ID: S0010101
     - Problems Solved: 87
   - Peers Selection
     - Frequency
     - Proportion
   - Problems

5. Interaction
Evaluation

RQ1: Is peer data useful?
RQ2: Is visualizing more suggestions helpful for planning compared to only one path?
RQ3: Does visualizing more suggestions using the proposed visualizations increase complexity?
Evaluation: Controlled User Study

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
## Evaluation: Questionnaires

<table>
<thead>
<tr>
<th>Informativeness</th>
<th>Questionnaire Statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>The information needed to plan a learning path is easy to access.</td>
</tr>
<tr>
<td>Q2</td>
<td>The information needed to plan a learning path is rich.</td>
</tr>
<tr>
<td>Q3</td>
<td>The information is sufficient to plan a learning path.</td>
</tr>
<tr>
<td>Q4</td>
<td>The system was helpful for me to find a proper learning path for a specific learning scenario.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decision making</th>
<th>Questionnaire Statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q5</td>
<td>I am confident that I find a suitable learning path for the learning scenario.</td>
</tr>
<tr>
<td>Q6</td>
<td>The system helps make adjustment according to previous performance.</td>
</tr>
<tr>
<td>Q7</td>
<td>The learning path design is intuitive.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Visual design</th>
<th>Questionnaire Statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q8</td>
<td>The learning path design helps me understand the suggested path.</td>
</tr>
<tr>
<td>Q9</td>
<td>It was easy to learn the system.</td>
</tr>
<tr>
<td>Q10</td>
<td>It was easy to use the system.</td>
</tr>
<tr>
<td>Q11</td>
<td>I would like to recommend this system to others.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>System Usability</th>
<th>Questionnaire Statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q12</td>
<td></td>
</tr>
</tbody>
</table>
Results: Using peer data is useful (RQ1)
Results: Visualizing more suggestions is useful (RQ2)
Results: Visualizing more suggestions using the proposed visualizations did not increase the complexity (RQ3)
How data can be used for learners to plan, **perform**, and reflect on their learning?
Persua: A Visual Interactive System to Enhance the Persuasiveness of Arguments in Online Discussion

Meng Xia, Qian Zhu, Xingbo Wang, Fei Nie, Huamin Qu, Xiaojuan Ma

CSCW 2022
Background

- Real-time adaptive feedback
- **Critical thinking skills, e.g., arguments writing**, a major element of learning framework 2030 by the Organization for Economic Co-operation and Development (OECD)

Taxonomy of Argument Components

Classical persuasive strategies of Aristotle, and Rapp, 2002 and Carlile et al., 2018


Semi-structured interviews with five domain experts on debating competitions
Here’s too many people already. The earth deserves better. I don’t want to pass on my depression and anxiety. Life sucks. Now do you see?

I was spanked as a kid. I really have no intention of making spanking the forefront of discipline since it really doesn’t accomplish much, but it does accomplish one thing: it teaches children that getting hit hurts. Why does this matter? Because kids often resolve to hitting as a way to problem-solve. I hit my sister all the time when I was younger, because she was a brat. My sister went through a bad biting phase. Both of these things occurred when we were all little, where you aren’t really aware of how other people feel. There’s no immediate inclination to put yourself in their shoes. That’s just something you grow into and (ideally) gets augmented by how you’re raised. So, one time, I hit my sister and my mom spanked me. She asked me if it felt good. I said “No,” and the light went off on my head as if to say “Holy shi...” is that what I’ve been doing to people.” Same with my sister. One day, my mom bit my sister’s finger just enough and, after that, my sister stopped biting me. Anyway, even if it’s wholly wrong, that still isn’t Stockholm Syndrome. There is a cycle in domestic violence and if you include being in that category then that’s really all it is: a cycle. There are a lot of abusers out there who know damn well what they’re doing is unacceptable, apologize, and then do it again. That’s part of the problem, since so many people continuously accept the apology.

Replier name: moonflower
Are you wanting to understand how other people can find it rewarding, or are you wanting to change your view on the world?
Case Study

Before:

Case Study

Before:

[Claim] Here’s too many people already. [Pathos] The earth deserves better. [Pathos] I don’t want to pass on my depression and anxiety. [Pathos] Life sucks. [Pathos] Now do you see?

After:

[Claim] There is no need to have children. [Logos] [Evidence] The world population is dramatically increasing, leading to more consumption of the natural resources. [Pathos] [Evidence] In addition, I saw my uncle’s children fight with each other. [Logos] [Evidence] Some people even abuse children. [Claim] Thus, I disagree with having children.
Planning

Performing
Data

Reflecting

Understanding and Analyzing

Redesigning

Educators (design loop)

Learners (learning loop)

How data can be used for educators to improve the design of learning materials?

Peerlens (CHI 2019)
Persua (CSCW 2022)

QLens (TVCG 2021)
StuGPTVis (TVCG 2024)
QLens: Visual Analytics of Multi-step Problem-solving Behaviors for Improving Question Design

Meng Xia, Reshika Palaniyappan Velumani, Yong Wang, Huamin Qu, Xiaojuan Ma

TVCG 2021
A Multi-step Problem

Five people stand in a line.


Move each person to their place in the line.
Mark stands ahead of Paul.
Helen stands ahead of Jane.
Paul stands behind Helen but ahead of Luke.
No boy is next to another boy in the line.

Consider the condition with the most restrictions: “No boy is next to another boy in the line.”
There are only 3 boys and 2 girls, so we have

```
Back  Boy  Girl  Boy  Girl  Boy  Front
```

“Mark stands ahead of Paul”, “Paul stands ahead of Luke”. Therefore,

```
Back  Luke  Girl  Paul  Girl  Mark  Front
```

“Helen stands ahead of Jane”. Therefore,

```
Back  Luke  Jane  Paul  Helen  Mark  Front
```
Our collaborator, TrumpTech, uses QLens to improve questions design. The company serves more than 500 schools in Hong Kong.
StuGPTViz: A Visual Analytics Approach to Understand Student-ChatGPT Interactions

Under review

VIS 2024
Motivation

To empower students to fully leverage the capabilities of ChatGPT in educational scenarios, understanding students’ interaction patterns with ChatGPT is crucial for instructors.

You
what is data visualization?

ChatGPT
Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

Data visualization is important because it allows humans to comprehend complex datasets more easily, helping to identify correlations, anomalies, and trends that might not be apparent in raw data. It's widely used in various fields such as business, science, healthcare, finance, and more to communicate insights, make informed decisions, and discover new knowledge.
However, this endeavor is challenging due to \textbf{the absence of datasets} focused on student-ChatGPT conversations and the \textbf{complexities in identifying and analyzing the evolutorial interaction patterns} within conversations.
Context

- Graduate-level data visualization
- 48 students
- 10 exercise
This diagram illustrates the updated Visual Analytics Pipeline (named the knowledge generation model), which involves a series of operations, representations, and analytical stages. Using GPT's assistance, please identify and fill in the blank nodes which associated Analytical Stages (1 to 5) in the process.

Analytical Stages:
- Stage 1 is [Stage_A]
- Stage 2 is [Stage_B]
- Stage 3 is [Stage_C]
- Stage 4 is [Stage_D]
- Stage 5 is [Stage_E]

Student Answer:
- Finding: Action: Insight: Hypothesis: Knowledge

Raw Conversation:
- Do you know the diagram of updated visual analytics pipeline?
  - Yes, the updated visual analytics pipeline typically consists of the following stages: 1. Data acquisition and integration 2. Data cleaning and preprocessing 3. Data transformation and feature extraction 4. Data mining and analysis 5. Visualization and exploration 6. Knowledge representation and reasoning 7. Decision making and action. This pipeline is iterative and interactive, allowing analysts to refine their understanding of the data and adjust their analysis and visualization approaches as needed.
Enable learners and educators to find insights from learning data and use the insights to make decisions for achieving personalized online learning.
Future HAI in Education

- More Comprehensive
- More Actionable
- More Effective
- More Context-aware
More Comprehensive

- Peerlens (CHI 2019)
- Persua (CSCW 2022)
- AlgoSolve (CHI 2022)
- Distributed Tutorship (LAK 2022)
- SeqDynamics (EuroVIS 2020)
- QLens (TVCG 2021)
- SolutionVis (AIED 2023)
- Visual Analytics K-12 (VIS 2019, Best Poster Award)
- BlockLens (L@S 2022)
- Mobile MOOCs (CHI 2022, Best Paper Award)
- StuGPTVis (TVCG 2024)
More Actionable

QLens: Learning analytics

CTAT (cognitive tutor authoring tools): Learning design

Teacher-AI collaboration: data-driven optimization for learning materials
More Effective

Learner-AI collaboration: Long-term monitoring and accompany

Multiple Rounds Conversations/Interactive Knowledge Map

Other data
More Context-aware - **Immersive Online Learning**

**Learning in VR?**
- **Cinematography Education** on a Soundstage in VR (ongoing, poster, VR 2023)
- VR Story for **Awareness of Covid Spread Threats** (Zhu et al., IJHCI 2023)
- Towards an Understanding of Asymmetric **Collaborative Visualization on Problem-solving** (Tong et al., VR 2023)

**Learning in AR?**
- Exploring Interactions with Printed Data Visualizations in Augmented Reality (Tong et al., TVCG 2023, Honorable Mention Award)

Other opportunities, e.g., **on-the-go learning**, digital twin campus?
Future Education for HAI

- Fairness
- Accountability
- Transparency
- Social responsibility
Human-AI Interaction in other domains?

- Autonomous Driving
- Roburt Surgery
- Sports
Human-AI Interaction for Personalized Online Learning

Welcome to join our lab!

Personal page: [https://www.xiameng.org](https://www.xiameng.org) Email: mengxia@tamu.edu