UC San Diego

Beyond Pair Programming: How Experts and Novices Navigate the Knowledge Gap with Pair Data Analysis


Andrew H. Yang


Yiqing (Brenda) Li


Achintya Rai


Abhi Sathvika Goriparty


Sam Lau

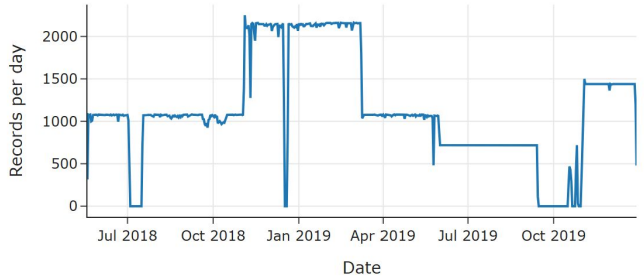
jupytercon **2025 JupyterCon | November, 2025**

Imagine...

Dataset	Platform	No. of Genes	No. of samples in class I	No. of samples in class II	Source
Colon	cDNA	2000	40(T)	22(N)	[18]
Leukemia	Affy	7129	25(AML)	47(ALL)	[19]
CNS	Affy	7129	25(C)	9(D)	[20]
DLBCL	Affy	7129	58(D)	19(F)	[21]
Lung	Affy	12533	150(A)	31(M)	[22]
Prostate1	Affy	12600	52(T)	50(N)	[23]
Prostate2	Affy	12625	38(T)	50(N)	[24]
Prostate3	Affy	12626	24(T)	9(N)	[25]
GCM	Affy	16063	190(C)	90(N)	[26]

CHROM	POS	ID	REF	ALT
chr1	871334	rs4072383	G	T
chr1	880238	rs3748592	A	G
chr1	881627	rs2272757	G	A
chr1	883625	rs4970378	A	G
chr1	887560	rs3748595	A	C

```
px.line(per_day, x=per_day.index, y='records_per_day',
        labels={'timestamp':'Date', 'records_per_day':'Records per day'},
        width=550, height=250,)
```



You find yourself asking...

What’s the best way to handle missing values in this context?

How do I visualize this data?

How do I interpret this pattern I’m seeing?

The “Classroom-to-Practice” Gap

- Data science students excel at **syntax** but often **struggle** when faced with **"messy," unstructured, real-world problems**
- The primary hurdles aren’t just coding; they’re **conceptual** (“what does this data mean?) and **strategic** (“What is the next step?”)

[1] Iraklis Varlamis. (2025). Messy Data in Education: Enhancing Data Science Literacy Through Real-World Datasets in a Master’s Program
[2] Singh, A., Fariha, A., Brooks, C., Soares, G., Henley, A. Z., Tiwari, A., Chethan M, Choi, H., & Sumit Gulwani. (2024). Investigating Student Mistakes in Introductory Data Science Programming

The “Obvious” Solution & Its Limits

- The natural solution seems to be AI assistants (e.g., Gemini in Sheets, Copilots)
- **What they do well:** Act as powerful **“task-completers”**. They instantly generate code, clean columns, and answer *specific* questions
- **Where they fall short:** They are **“answer-givers”**, not **“mentors”**. This approach risks promoting **superficial learning** and **over-reliance**

[1] Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025, June 10). *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task*.
[2] Delikoura, I., Fung, Yi. R., & Hui, P. (2025). *From Superficial Outputs to Superficial Learning: Risks of Large Language Models in Education*. ArXiv.org. https://www.arxiv.org/abs/2509.21972

Therefore Our Goal was to:

Build a tool, embedded directly into the **data science workflow (JupyterLab)**, to provide **guided mentorship**—not just answers.

Watchful Intelligent Science Expert

WISE for short

What is WISE?

AI Assistant

Real Time & Convenient

Domain Aware Support

Relevant Insights within Jupyter Environment

Smarter EDA Suggestions

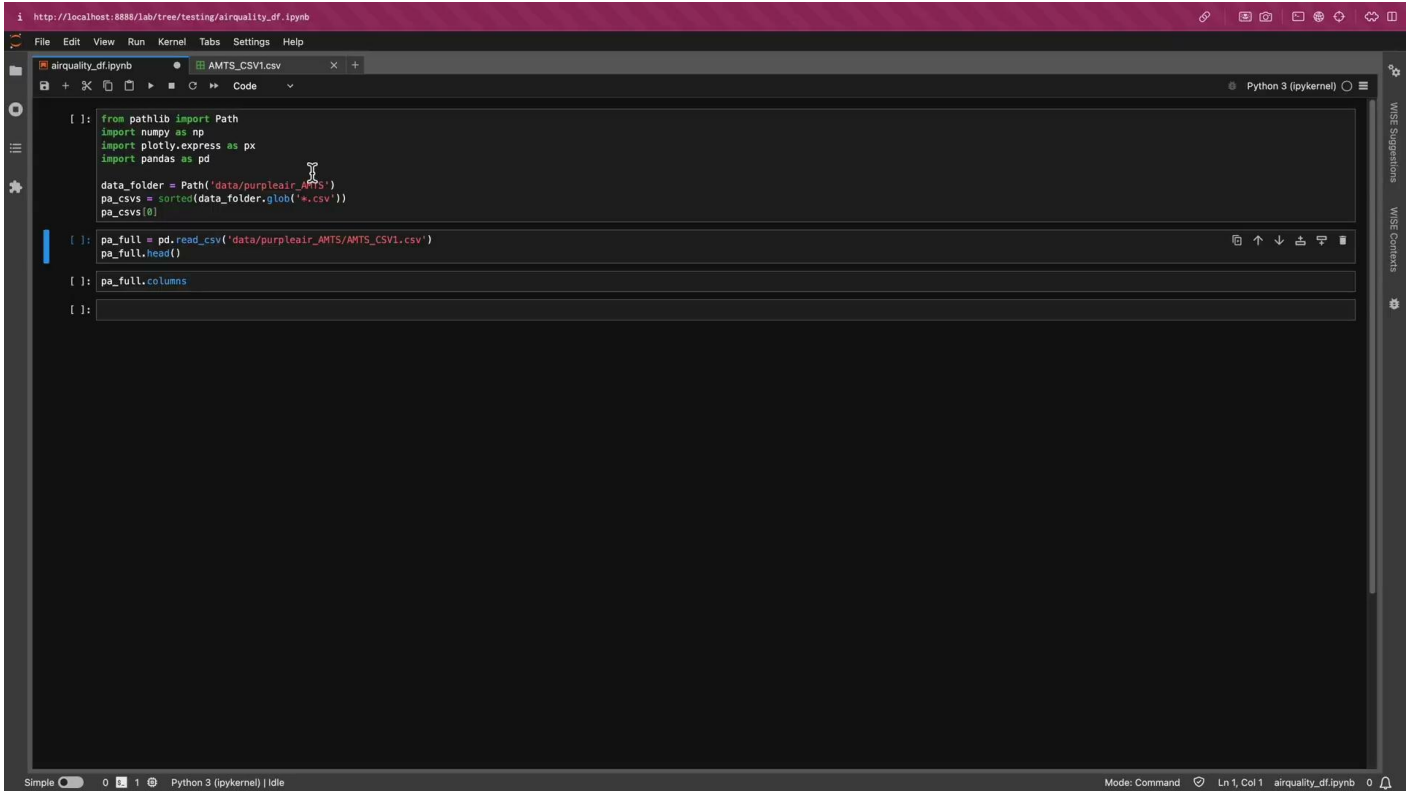
Task-Specific Guidance

How did we design it?

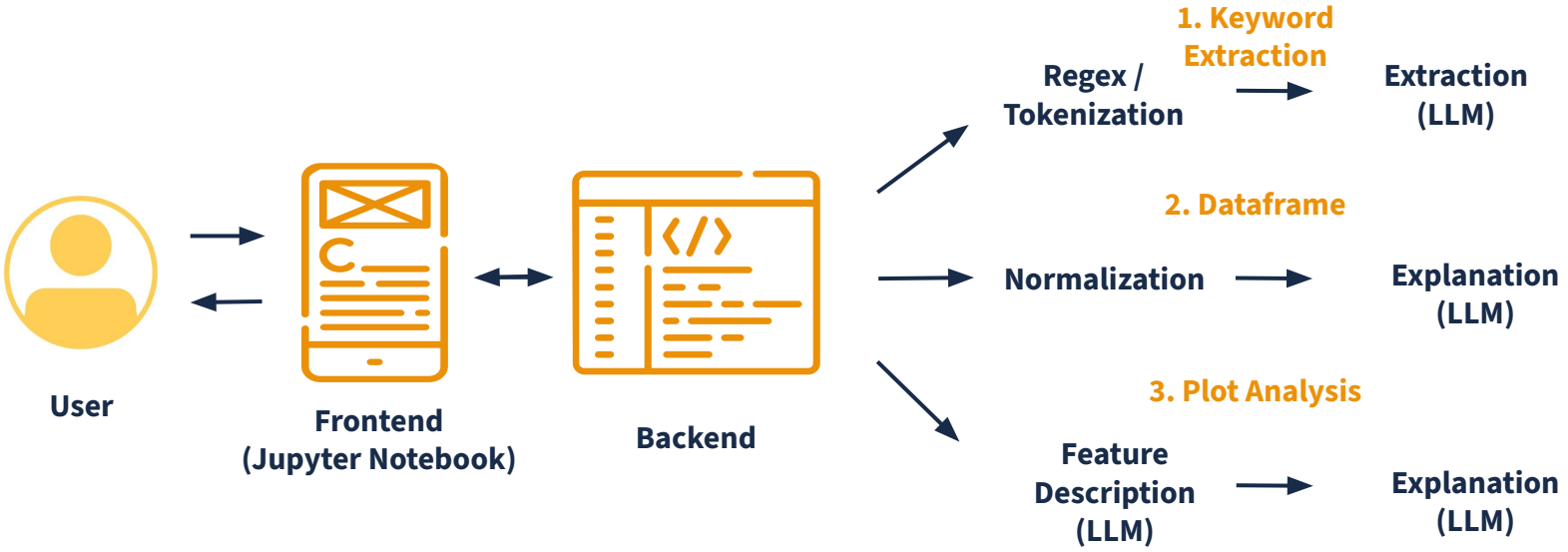
We focused on design on these principles:

- **Ease of Use** - Proactively assists users
- **Data-Specific Analysis** - Features target key analysis bottlenecks (e.g., parsing, terminology, graph trends)
- **Flexible Architecture** - Built for scalability, maintainability, and easy addition of new features

WISE Demonstration



Technical Architecture



Technical Architecture (Cond.)

Modularity	• Easily add new features (plot analysis, dataframe analysis) • Aids in concurrent development
Decoupled Architecture	• Easily refactor code for UI/UX Changes • Standardized data format
Reusable Context	• Extract keywords once → use multiple times • Increase response quality without large overhead
Fact Verification	• Multi-stage LLM processing • Significantly reduce hallucinations

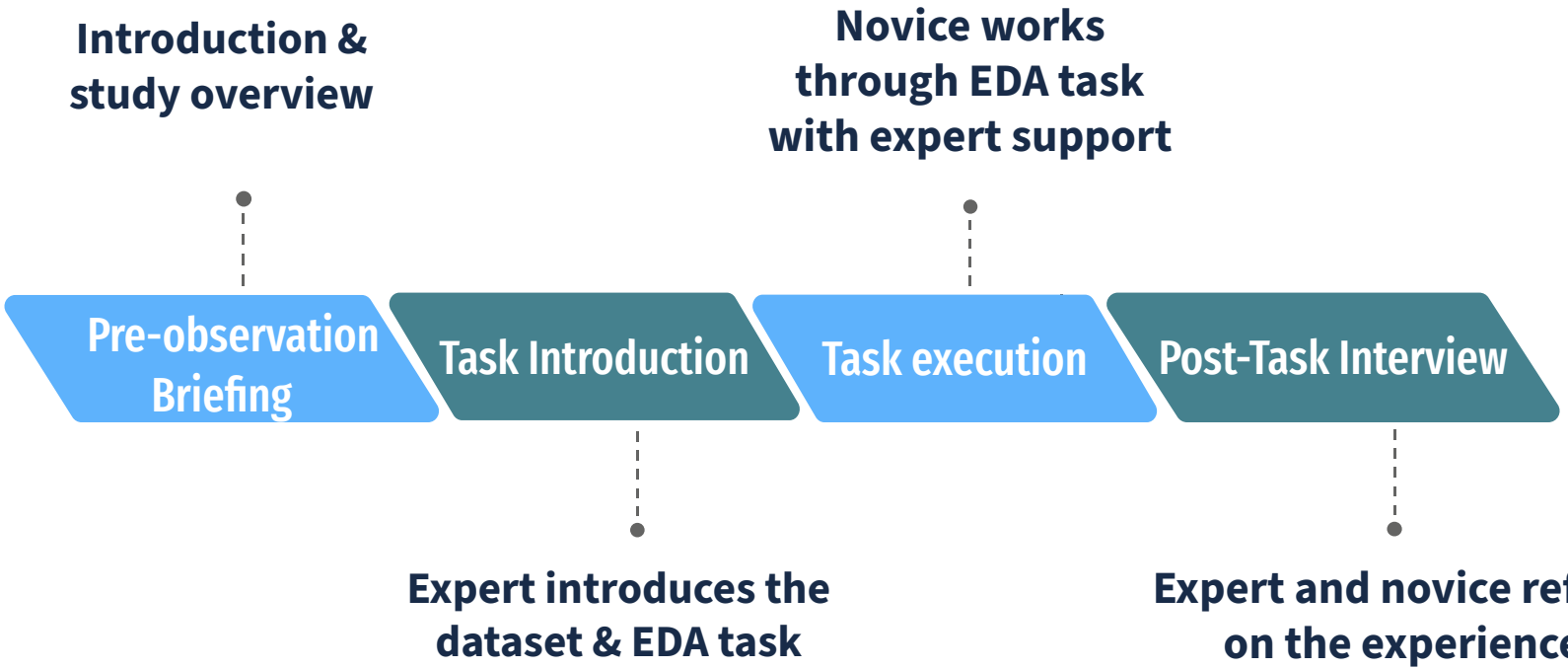
However, we're not sure how effective WISE's support will be for novices.

RQ: How do experts help novice data scientists make sense of and work effectively with data from unfamiliar domains?

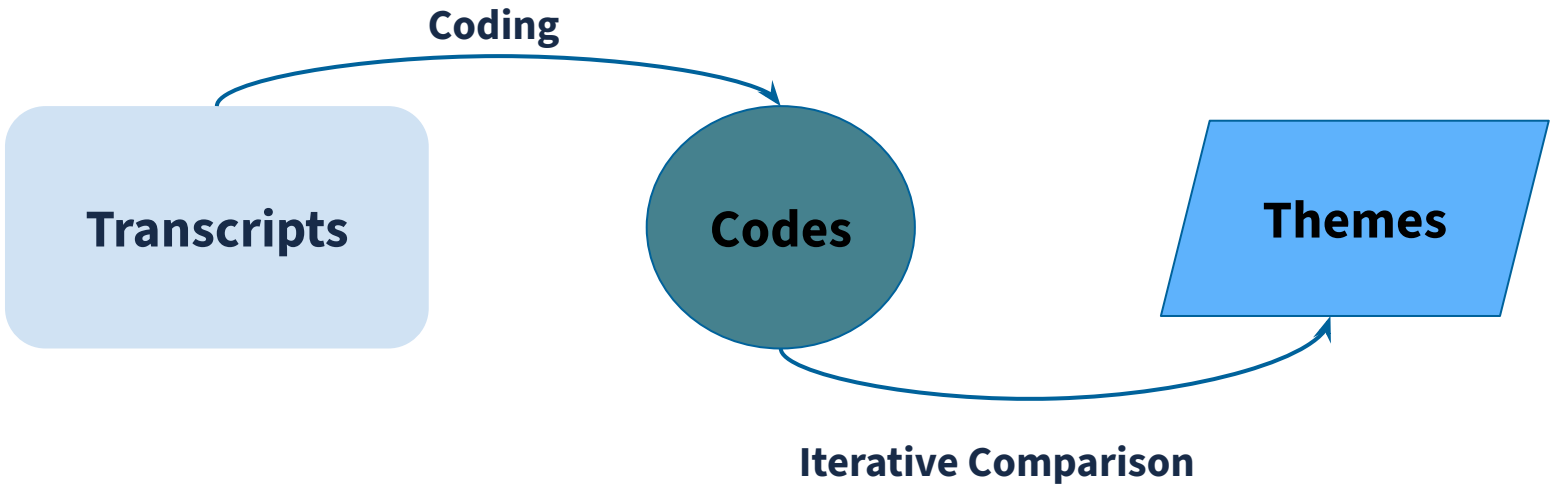
We Paired 7 Domain Experts with 7 Novice Data Scientists

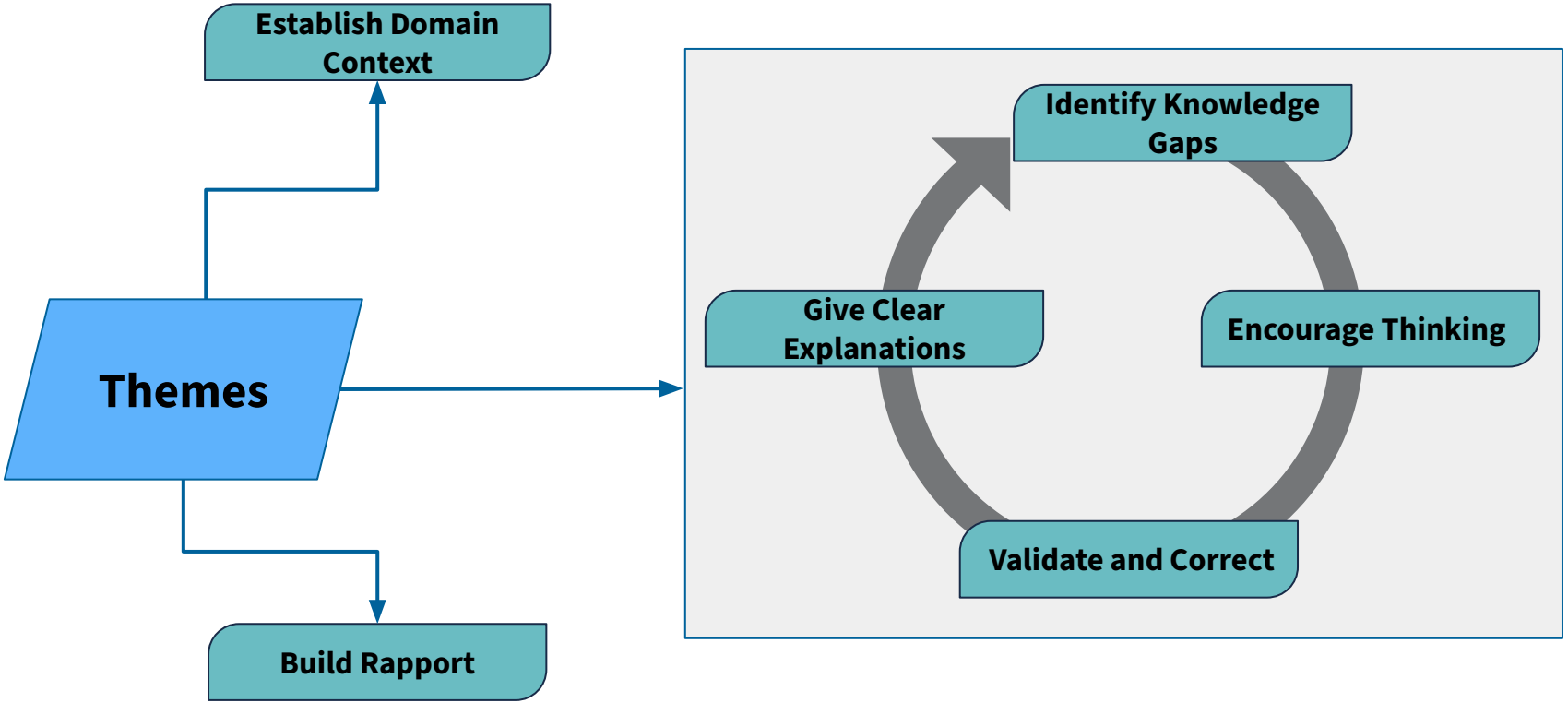
- Domain Experts
  - PhD researchers in diverse domains (e.g., data privacy, climate)
  - Designed domain-specific EDA tasks in Python
- Novice Data Scientists
  - 3rd-year DS undergrads
  - Unfamiliar with the domain and dataset they were assigned

We Observed How Experts and Novices Collaborate



Interpreted Transcripts Using Thematic Analysis

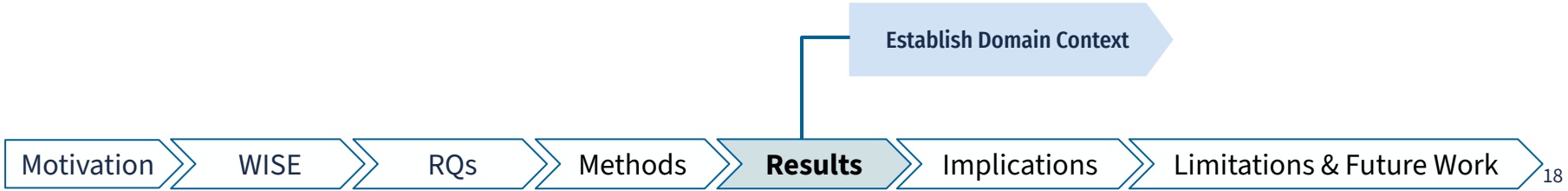



Establish Domain Context
Themes
Build Rapport
Identify Knowledge Gaps
Encourage Thinking
Validate and Correct
Give Clear Explanations

How Experts Prepare Novices to Work with Unfamiliar Data:

Helps novices make sense of data through real-world examples:

"It seems that whenever you have a low night, you seem to have a high night later… that's because of something called sleep pressure. ***If you lose a lot of sleep one night, your body will fall asleep quicker and spend more time in deep and REM stages.****"*

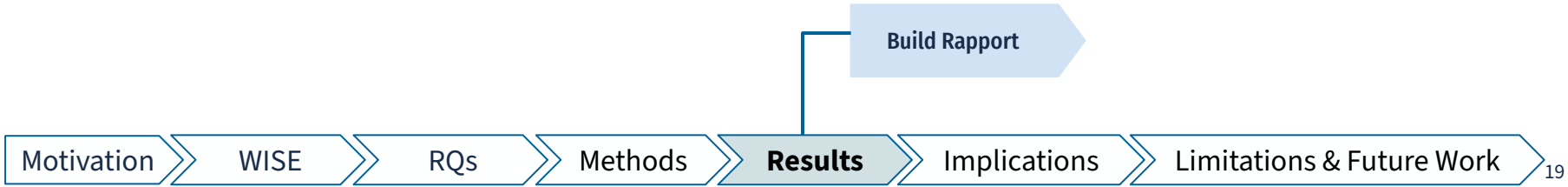


How Expert Build Rapport with Novices:

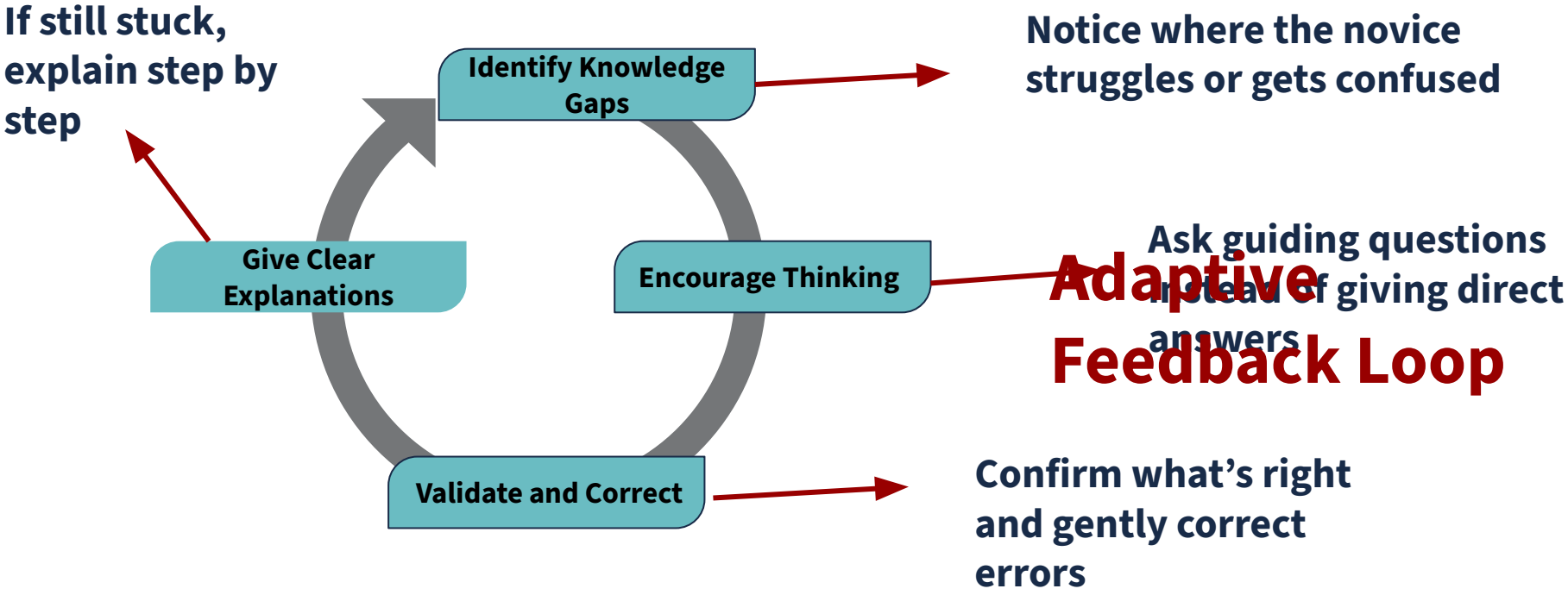
Reassurance - to reduce anxiety:

“That’s okay, we’ll just work through it as a challenge.”

“I think that's very close to the expected output. You may just rename these columns.”



How Experts Help Novices Make Progress:



Implications – AI Tutors

- Adaptive support:
  - tailor the **amount** and **type** of support to learner needs
  - balance guidance and independence
- Teach **analytical reasoning**, not just answers:
  - encourage learners to explain and reflect on their thinking
  - develop transferable reasoning skills

Implications – Educators

- Build **contextual understanding**:
  - connect data to real-world relevance
- Use adaptive scaffolding:
  - support learning through **adaptive feedback loops**
- Build rapport and **motivation**:
  - provide encouragement, validation, and reassurance

Limitations & Challenges

- Moments of **silence**:
  - Preemptively identify when novice is stuck
- Inconsistent EDA tasks:
  - Hard to compare effectiveness across different task formats and domains

Future Work

- Conduct a **large-scale quantitative study**:
  - Compare performance with and without WISE
- Make WISE more **adaptive and dynamic**:
  - Adjust the type and amount of help based on users' needs

Questions?

UC San Diego

- **The problem:**
  - Novices struggle with **analysis strategy** ("what to do next") and **domain knowledge gaps** .
- **Our Tool (WISE):**
  - A proactive JupyterLab extension built to be a **guiding mentor** .
- **Our Study:**
  - We analyzed human expert-novice pairs to define **effective mentorship** .
- **Key Findings:**
  - Experts adapt: they **guide strategy** with questions and **build confidence** with validation.

UC San Diego

Thank you!

For More Questions:

Contact Andrew H. Yang: any012@ucsd.edu or


Andrew H. Yang


Yiqing (Brenda) Li


Achintya Rai


Abhi Sathvika Goriparty


Sam Lau

UC San Diego

SCHOOL OF COMPUTING, INFORMATION AND DATA SCIENCES

HALICIOĞLU DATA SCIENCE INSTITUTE