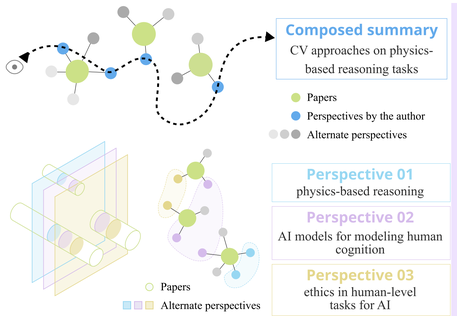
PersLEARN: Research Training through the Lens of Perspective Cultivation

PersLab Research

61 ACL 2023

On Perspective Cultivation

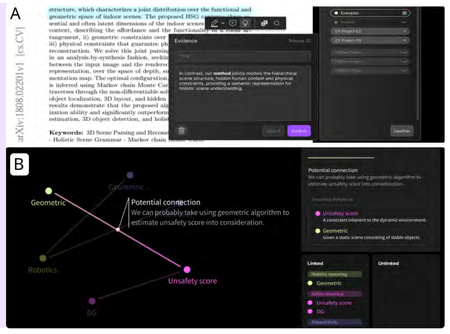


Composed summary vs. framed perspectives

Composed summaries are subject to the authors' perspectives, whereas the perspective frames are directed by new ideas.

PersLEARN UI showcases

(A) A selected piece of evidence and its interpretation. (B) A generated perspective frame.



Taking-away Messages

Problem: Junior researchers often face challenges in identifying the perspectives reflected in the existing literature and struggle to develop their own viewpoints.

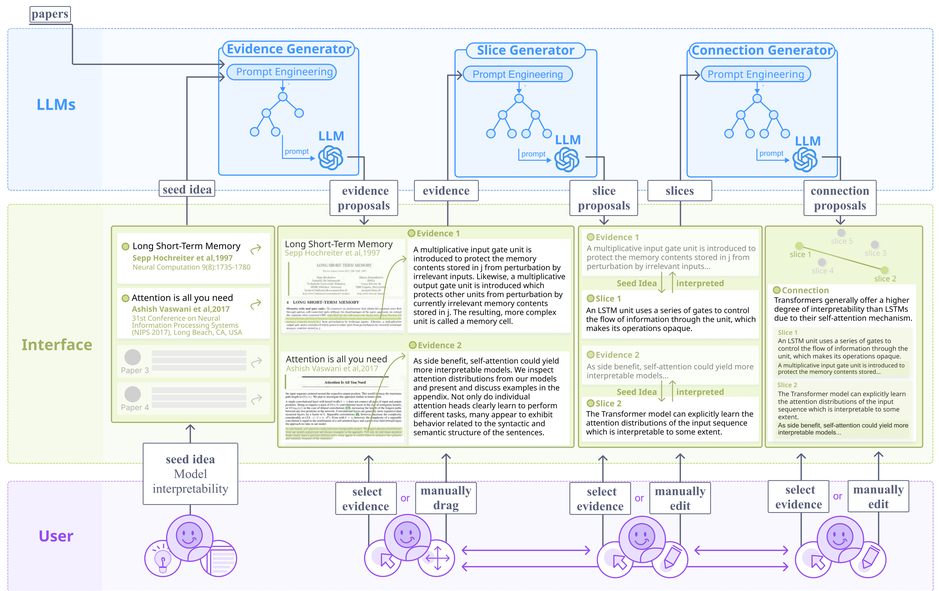
Solution: PersLEARN is a tool designed to facilitate the cultivation of scientific perspectives, starting from a basic seed idea and progressing to a well-articulated framework.

Method: Researchers develop their perspectives explicitly by interacting with an LLM.

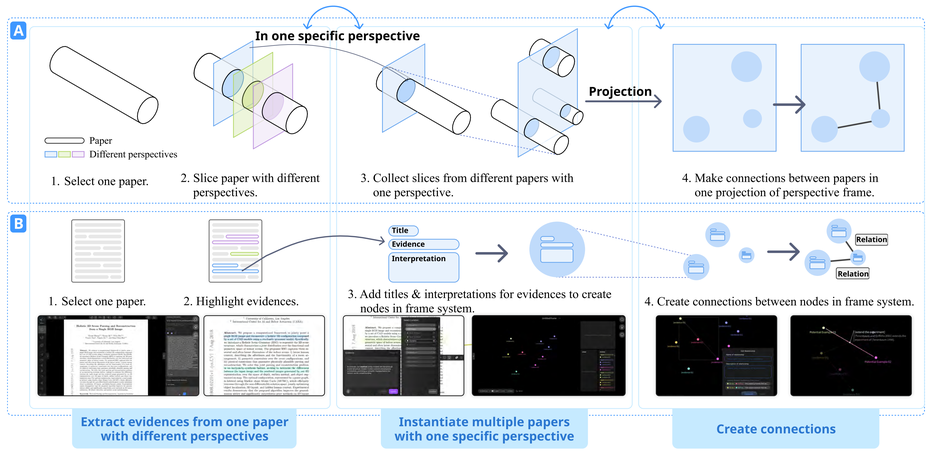
Result: Students using PersLEARN exhibit a superior level of logical coherence and depth compared to those that did not.

Significance: Our framework of scientific perspective may bring science education to a future with better student-centered considerations.

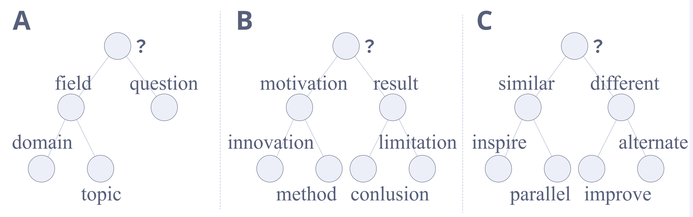
System Diagram



Theoretical Framework & User Interface



Semi-automation & evaluation



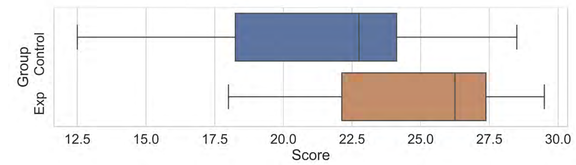
Grammar for Prompt Engineering (A) Evidence location. (B) Slice generation. (C) Connection generation.

Table 1: **Result of our pipeline.** w/o and w/ are with and without prompt engineering, respectively. The performance of slice generation, connection generation, and perspective diversity indicate the efficacy of our prompt engineering.

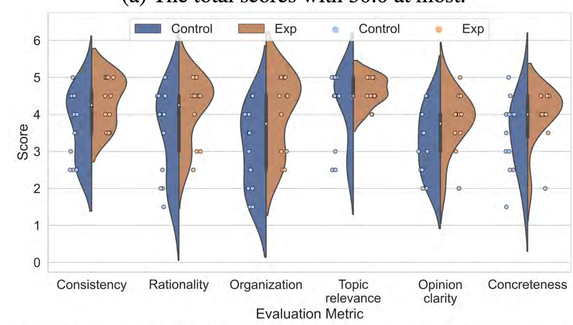
Metric	w/o	w/
Slice BLEURT	0.238	**0.795**
Connection CR	0.450	**0.550**
Perspective VMR	0.028	**0.006**

36 perspectives, 333 papers in the literature

Human Study



(a) The total scores with 30.0 at most.



(b) The scores of the 6 metrics with 5.0 for each. The presented scores in the plot may surpass 5.0 due to smoothing.

Seed idea: Physics-based reasoning in Computer Vision.

1 Holistic 3d Scene Parsing and Reconstruction from a Single RGB Image. ECCV, 2018.

2 Scene Understanding by Reasoning Stability and Safety. IJCV, 2015.

3 Galileo: Perceiving Physical Object Properties by Integrating a Physics Engine with Deep Learning. NeurIPS, 2015.

4 Physics 101: Learning Physical Object Properties from Unlabeled Videos. BMVC, 2016.

5 Understanding Tools: Task-oriented Object Modeling, Learning, and Recognition. CVPR, 2015.

The complexity and richness of human vision are not only reflected by the ability to recognize visible objects, but also to reason about the latent actionable information, including inferring physical object properties[4] or latent human context as the functionality of a scene[1]. But how does human vision achieve such complexity and richness ? Some studies propose a model based on physics principle for understanding these mechanisms[2,3,4]. Even other sudies consider more complex situations, where humans have tasks[5] and other hidden human context[1].

w/o PersLEARN

Physics-based reasoning has been used for two aims. The first is to learn the physical properties of an object. For example, Galileo[3] and Physics 101[4] learn physical properties like mass and density from videos. Bo Zheng et al[2] learn stability and safety of objects in a scene. The second is to enrich the object representation by incorporating physical features. The enriched representation is then used to assist other visual tasks. Siyuan Huang et al[1] design a physically enriched HSG representation of 3D scene structure in the single-view 3D reconstruction task. Yixin Zhu et al[5] use a representation consisting four physical-functional components in object recognition task.

w/ PersLEARN