CSCE 689 - Special Topics in NLP for Science (Spring 2025)

Lecture 1: Overview

Yu Zhang

yuzhang@tamu.edu

January 14, 2025

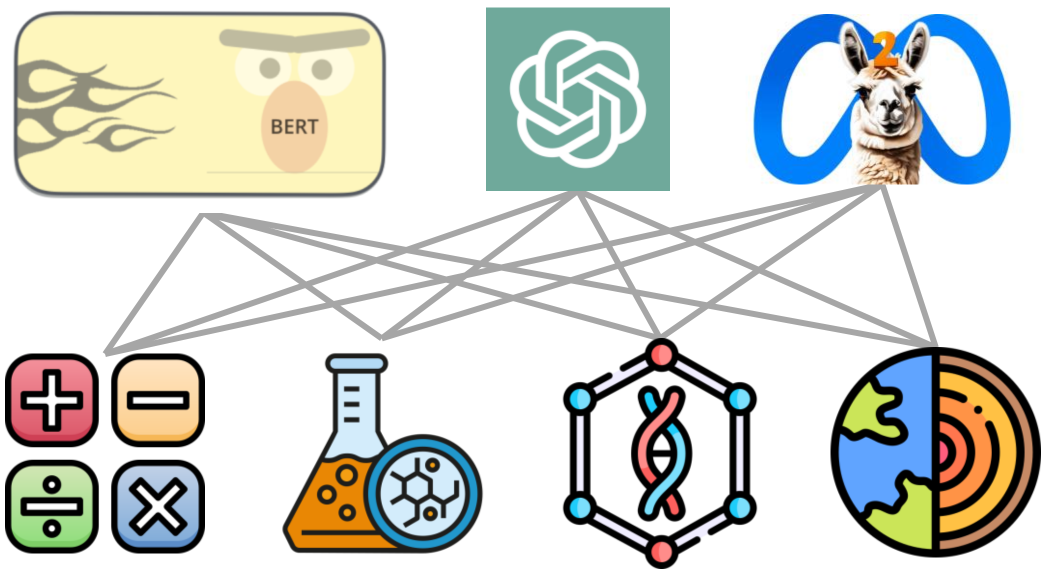
Course Website: https://yuzhang-teaching.github.io/CSCE689-S25.html

Good News (Or Bad News?)

- No midterm or final exam
- No more classical NLP content

What an NLP course used to be … → What this NLP course will be …

Sep 30 Lecture 3: Machine Translation: Word Alignment, Parallel Corpora, Decoding, Evaluation	
Oct 7 Lecture 5: N-Grams, Final Project Discussion	
Oct 14 Lecture 7: Competitive Grammar Writing I	
Oct 21 Lecture 9: Dependency Parsing	Final project proposal due
Oct 28 Lecture 11: Coreference Resolution II	



Course Logistics

- Instructor: Yu Zhang (yuzhang@tamu.edu)
- Lectures:
    - Time: Tuesdays and Thursdays 3:55pm – 5:10pm
    - Location: HRBB 126
- Office Hour:
    - Time: Thursdays 2pm – 3pm
    - Location: PETR 222 (or drop me an email at least 1 day in advance if you would like to join via Zoom: https://tamu.zoom.us/j/6411788612)

Course Format & Grading

- Course website: https://yuzhang-teaching.github.io/CSCE689-S25.html

- Paper presentation (20%)
  - Each student will sign up for a lecture and present the corresponding 3 papers.

- Only purple rows on the schedule can be selected.
- Most of them are on Thursdays.

	2/6	Paper Classification Can be selected	Papers to be presented * The Effect of Metadata on Scientific Literature Tagging: A Cross-Field Cross-Model Study [WWW 2023] * Hierarchical Multi-Label Classification of Scientific Documents [EMNLP 2022] * BERTMeSH: Deep Contextual Representation Learning for Large-Scale High-Performance MeSH Indexing with Full Text [Bioinformatics 2020]		Student
W5	2/11	Scientific VLMs: Bioimaging Cannot be selected	* MedCLIP: Contrastive Learning from Unpaired Medical Images and Text [EMNLP 2022] * A Visual–Language Foundation Model for Pathology Image Analysis using Medical Twitter [Nature Medicine 2023] * LLaVA-Med: Training a Large Language-and-Vision Assistant for Biomedicine in One Day [NeurIPS 2023] * A Generalist Vision-Language Foundation Model for Diverse Biomedical Tasks [Nature Medicine 2024]		Instructor
	2/13	Scientific VLMs: Geometry Can be selected	* UniMath: A Foundational and Multimodal Mathematical Reasoner [EMNLP 2023] * G-LLaVA: Solving Geometric Problem with Multi-Modal Large Language Model [arXiv 2023] * Math-LLaVA: Bootstrapping Mathematical Reasoning for Multimodal Large Language Models [EMNLP 2024]		Student
W6	2/18	[Guest Lecture] Hanwen Xu (University of Washington): Towards Patient Level Representations for Better Clinical Outcome * Suggested Reading: A Whole-Slide Foundation Model for Digital Pathology from Real-World Data [Nature 2024]			Guest Lecturer
	2/20	Scientific VLMs: Miscellaneous Can be selected	* UrbanCLIP: Learning Text-Enhanced Urban Region Profiling with Contrastive Language-Image Pretraining from the Web [WWW 2024] * BioCLIP: A Vision Foundation Model for the Tree of Life [CVPR 2024] * MMMU: A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI [CVPR 2024]		Student

Course Format & Grading

- Paper presentation (20%)
  - Each student will sign up for a lecture and present the corresponding 3 papers.
  - I will send an email to everyone when the sign-up sheet is open (later today).
  - Slots are first come, first served!

  - The earliest two slots are 1/30 and 2/6 (Thursdays on Week 3 and Week 4).
  - Students selecting these two slots will be given 2% and 1% extra credit, respectively.

Course Format & Grading

- Paper presentation (20%)
  - Email your slides to the instructor at least 2 days before your presentation.
    - For a Thursday lecture, the slides should be emailed by Tuesday 11:59pm.
    - Start preparing your presentation early (e.g., 10+ days in advance).
  - Presentation duration: Strictly limited to 60 minutes, followed by a 15-minute Q&A session with the audience.

Course Format & Grading

- Paper presentation (20%)
  - Rubrics:
    - Slides (5%): Email your complete slides before the deadline.
    - Presentation Completeness (5%): Adequately cover the core concepts and insights contained in the paper. Feel free to emphasize more on the intuition and omit overly detailed parts.
    - Presentation Clarity (5%): Explain your own understandings of the papers in your presentation (e.g., raise a new example when introducing some concepts, list some limitations not mentioned in the paper).
    - Question and Answering (5%): Effectively answer the questions raised by the audience.

Course Format & Grading

- Project (60%)
  - Complete a research project, present your results, and submit a project report.
  - Work in a team of 2 or 3 (any deviation from this size requires prior approval from the instructor).
  - Possible topics:
    - Type I, survey: Carefully examine and summarize existing literature on a topic covered in this course; provide detailed and insightful discussions on the unresolved issues, challenges, and potential future opportunities within the chosen topic.

Course Format & Grading

- Project (60%)
  - Possible topics:
    - Type I, survey
    - Type II, hands-on project
      - Develop an effective algorithm for a scientific NLP task, or
      - Build a novel benchmark dataset for a scientific NLP task, or
      - Analyze the behavior of some existing scientific language models, or
      - …

      - Related to a topic covered in this course
      - Eligible for extra credits if publishable

Course Format & Grading

- Project (60%)
  - Project proposal due: 2/16 (Sunday) — 5%
  - Midterm spotlight presentation: 3/20 (Thursday) — 5%
  - Midterm report due: 3/23 (Sunday) — 10%
  - Final project presentation: 4/22 (Tuesday) and 4/24 (Thursday) — 15%
  - Final report due: 5/4 (Sunday) — 25%

Course Format & Grading

- Literature Review (10%)
  - Submit a review for a paper introduced in the lectures.
    - You can choose any paper on the schedule except the papers presented by you in your lecture.
  - The review should include a paper summary, strengths, weaknesses, questions to the authors, and limitations.
  - Due: 3/7 (Friday)

  - You cannot use large language models to help you write the review (except for grammar check).
  - You cannot copy from publicly available reviews of the paper.

Summary:

The authors fine-tune large PLMs (pre-trained on a corpus of scientific and mathematical text) on a new quantitative reasoning dataset, evaluating on existing benchmarks and a new, custom benchmark. The main contribution is the formulation of these datasets, which allows for sota on MATH and strong performance on other benchmark. Notably, this approach relies exclusively on a language model, which has no outside access to libraries, calculators, or other tools in performing the task.

Strengths And Weaknesses:

Strengths

- The most obvious strength of the paper is the considerable improvement over SOTA scores, although I'm quite surprised at how low such scores are, especially when GPT-3 also so exceeds them.
- I think the majority voting idea is really cool, since sampling obviously provides a limited view of the branching probability tree characterizing the language model's processing of the original prompt, and majority rule arguably offers a more fulsome picture of the outputs. The obvious problem with it is that cost scales linearly.
- Checking the answers using sympy is a clever way to bin different expressions of answers.

Weaknesses

- The main problem with the paper, in my view, is that the training data is the main contribution of the paper, but I cannot tell what it is like. Basically the only thing I can tell is that it contains mathematical expressions. Fig. 1 and Fig. 2 show the evaluation data, which is already outlined in other work, and I wonder why that space wasn't used to explain this paper's training corpus. At one point, the authors mention "equations and diagrams" juxtaposed with natural language, but how would a language model process diagrams, or how would they be converted to text? I have many more questions, and they might be answered in the appendix, which wasn't available to me, but this lack of clarity seems to me to preclude the publication of this paper at NeurIPS.
- The benchmark introduced may be too small to be meaningful, and I don't think the paper makes it clear how representative it is. Furthermore, I worry, when it includes questions from "'solid-state chemistry', 'information and entropy', 'differential equations', and 'special relativity'", that quantitative reasoning is not what we are testing, but specific scientific knowledge that makes use of quantitative reasoning.
- Doesn't reference EQUATE (Ravichander et al. 2019), one such recent, decently-cited quantitative reasoning benchmark.
- The authors say in L51 that "existing benchmarks are limited with respect to quantitative reasoning", and I'd like to know what they mean by that. I couldn't find any corresponding discussion.

Questions:

- The claim in L175-177 seems unfounded to me. Why wouldn't you be sure you haven't found a local minimum with the fine-tuned model by training on the wrong dataset, whereas you haven't with the base model?
- In L57-58, I cannot parse "Prompting language models... unseen problems". Do you mean to say "apply" instead of "output"?
- I cannot tell whether the pre-trained model, on which Minerva is based, is a contribution of this paper or another paper?
- I don't understand the need for the OCW benchmark introduced, in addition to MATH, GSM8k, and others. Which part of the new dataset do these more established benchmarks lack?

Limitations:

The authors rightly identify 3 limitations of their method. There is nothing set in stone, however, about these limitations (one could engineer a dataset to give specific capabilities to a model, or use a model that had access to outside tools, etc), and the authors do not discuss why they chose the model, data, etc. that they did **with respect to** these limitations.

Example: https://openreview.net/forum?id=IFXTZERXdM7&noteId=fWyUVKIcadp

Course Format & Grading

- Participation (10%)
  - Attendance: There are 28 lectures (including guest lectures, student lectures, and midterm/final project presentations) in total. You are required to attend at least 20 lectures.
  - If it is your turn to give a presentation (i.e., your lecture and your midterm/final project presentation), you cannot be absent.
  - Please refer to Student Rule 7 if you need to request exceptional absences.
  - Introduce yourself!

Course Format & Grading

- Participation (10%)
  - Pre-lecture Questions: Read the papers to be introduced in each student/guest lecture and submit a question you come up with.
    - The deadline is one day before the lecture (e.g., For Thursday lectures, you need to submit the question by Wednesday 11:59PM).
    - We will use Google Forms to collect pre-lecture questions.
    - You are required to submit at least 5 pre-lecture questions (at most 1 per lecture) throughout the semester.

- You only need to submit pre-lecture questions for student/guest lectures.

	2/6	Paper Classification You can submit	* The Effect of Metadata on Scientific Literature Tagging: A Cross-Field Cross-Model Study [WWW 2023] * Hierarchical Multi-Label Classification of Scientific Documents [EMNLP 2022] * BERTMeSH: Deep Contextual Representation Learning for Large-Scale High-Performance MeSH Indexing with Full Text [Bioinformatics 2020]		Student
W5	2/11	Scientific VLMs: Bioimaging No need to submit	* MedCLIP: Contrastive Learning from Unpaired Medical Images and Text [EMNLP 2022] * A Visual–Language Foundation Model for Pathology Image Analysis using Medical Twitter [Nature Medicine 2023] * LLaVA-Med: Training a Large Language-and-Vision Assistant for Biomedicine in One Day [NeurIPS 2023] * A Generalist Vision-Language Foundation Model for Diverse Biomedical Tasks [Nature Medicine 2024]		Instructor
	2/13	Scientific VLMs: Geometry You can submit	* UniMath: A Foundational and Multimodal Mathematical Reasoner [EMNLP 2023] * G-LLaVA: Solving Geometric Problem with Multi-Modal Large Language Model [arXiv 2023] * Math-LLaVA: Bootstrapping Mathematical Reasoning for Multimodal Large Language Models [EMNLP 2024]		Student
W6	2/18	You can submit	[Guest Lecture] Hanwen Xu (University of Washington): Towards Patient Level Representations for Better Clinical Outcome * Suggested Reading: A Whole-Slide Foundation Model for Digital Pathology from Real-World Data [Nature 2024]		Guest Lecturer
	2/20	Scientific VLMs: Miscellaneous You can submit	* UrbanCLIP: Learning Text-Enhanced Urban Region Profiling with Contrastive Language-Image Pretraining from the Web [WWW 2024] * BioCLIP: A Vision Foundation Model for the Tree of Life [CVPR 2024] * MMMU: A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI [CVPR 2024]		Student

Course Format & Grading: Summary

- Participation (10%)
  - (Attending 20 lectures + Submitting 5 questions) × 0.4% = 10%
- Literature Review (10%)
- Paper presentation (20%)
  - Slides (5%) + Presentation Completeness (5%) + Presentation Clarity (5%) + Question and Answering (5%)
- Project (60%)
  - Project Proposal (5%) + Midterm Presentation (5%) + Midterm Report (10%) + Final Presentation (15%) + Final Report (25%)

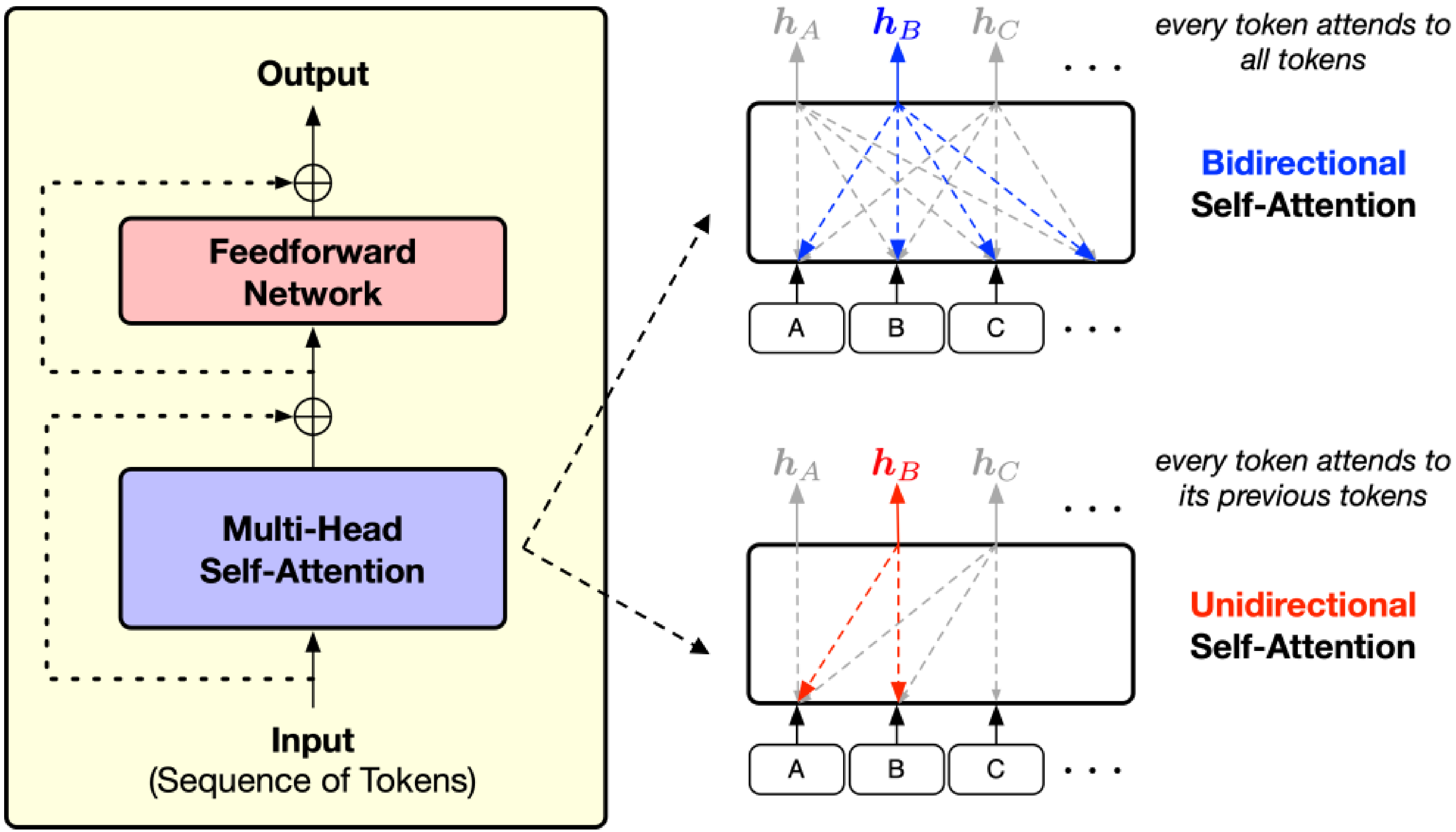
Overview of Course Contents

- Scientific Large Language Models
  - Encoder-Only, Decoder-Only, Encoder-Decoder
- Fundamental Scientific NLP Tasks
  - Citation Prediction, Literature Retrieval, Question Answering, Knowledge Extraction, Paper Classification
- Scientific Large Vision-Language Models
  - Bioimaging, Geometry, Geography, …
- Scientific Language Models for Other Data Modalities
  - Protein, DNA/RNA, Molecule, Academic Graph, Table, …
- Scientific NLP for Automating Research
  - Idea Generation, Content Generation, Execution, Reviewing, …

Overview of Course Contents

- Scientific Large Language Models
  - Encoder-Only, Decoder-Only, Encoder-Decoder
- Fundamental Scientific NLP Tasks
  - Citation Prediction, Literature Retrieval, Question Answering, Knowledge Extraction, Paper Classification
- Scientific Large Vision-Language Models
  - Bioimaging, Geometry, Geography, …
- Scientific Language Models for Other Data Modalities
  - Protein, DNA/RNA, Molecule, Academic Graph, Table, …
- Scientific NLP for Automating Research
  - Idea Generation, Content Generation, Execution, Reviewing, …

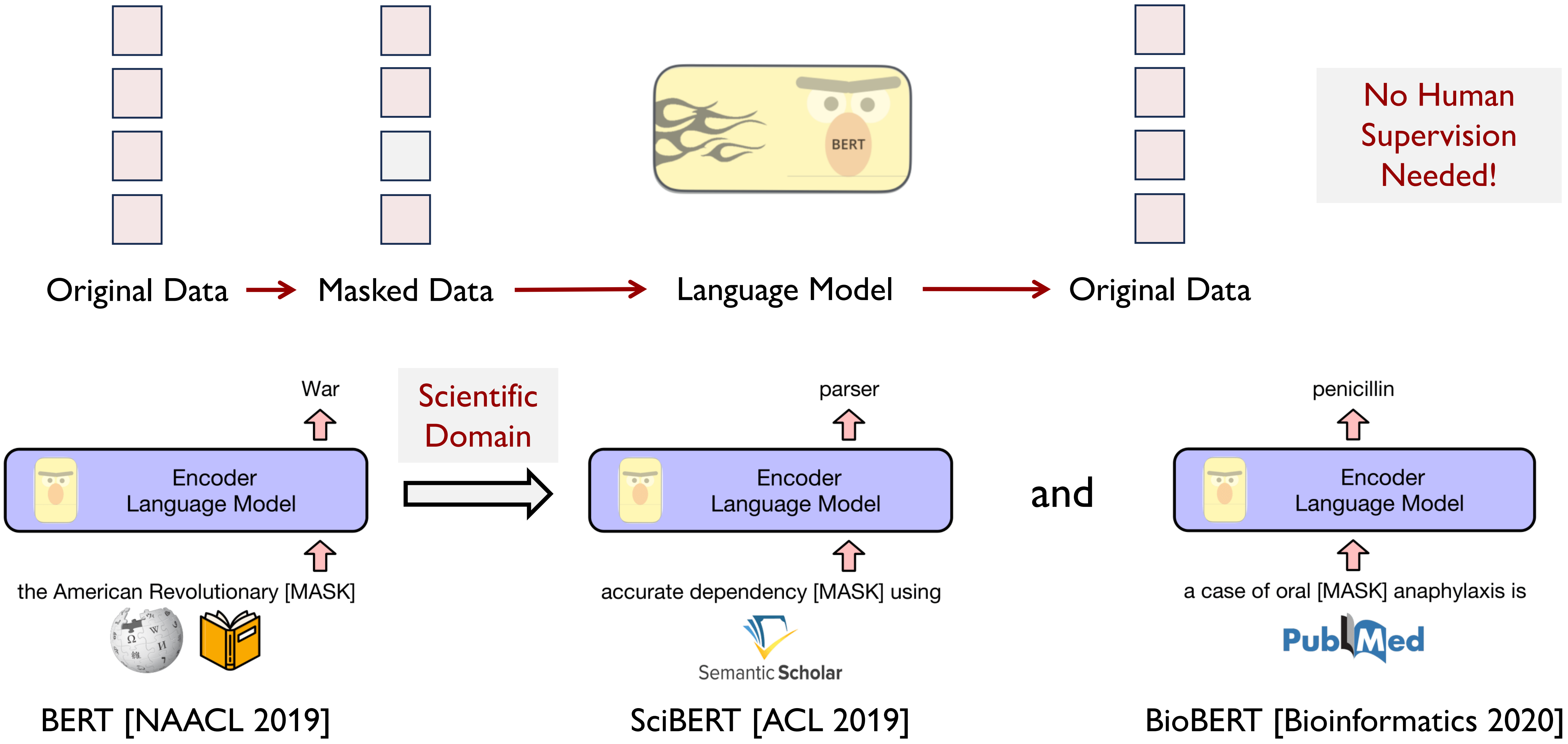
The Transformer Architecture



Transformer Encoder (Understanding)

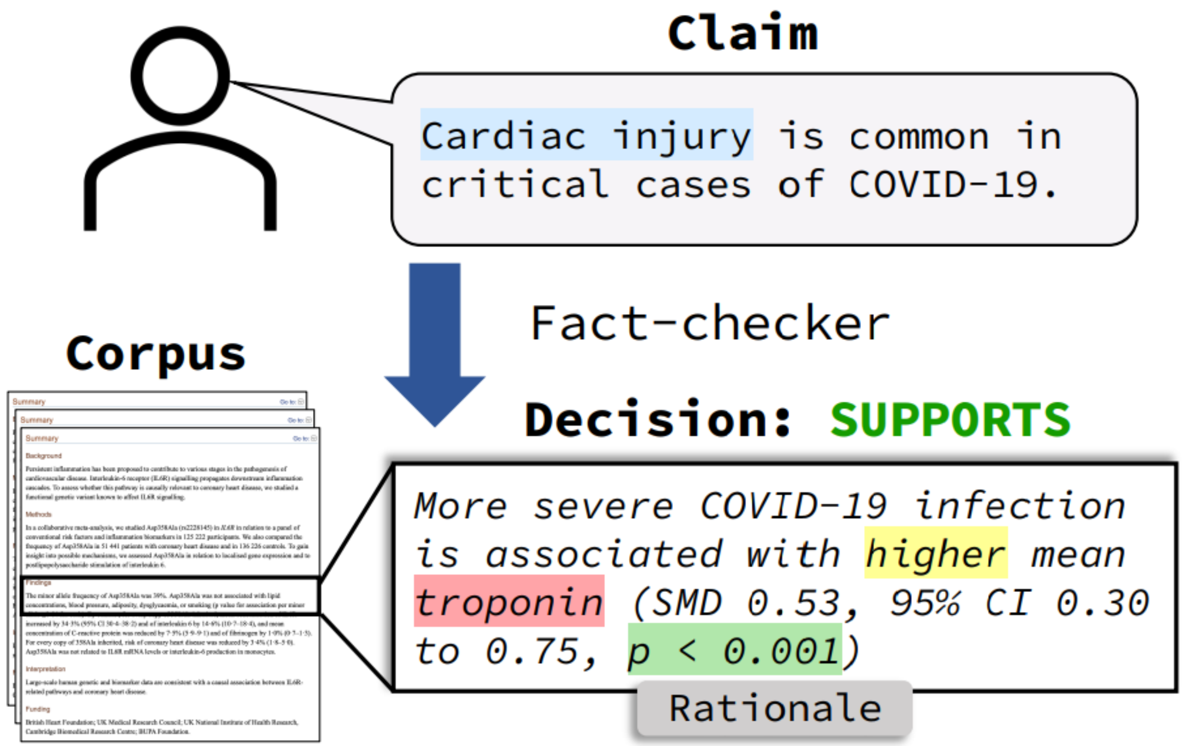
Transformer Decoder (Generation)

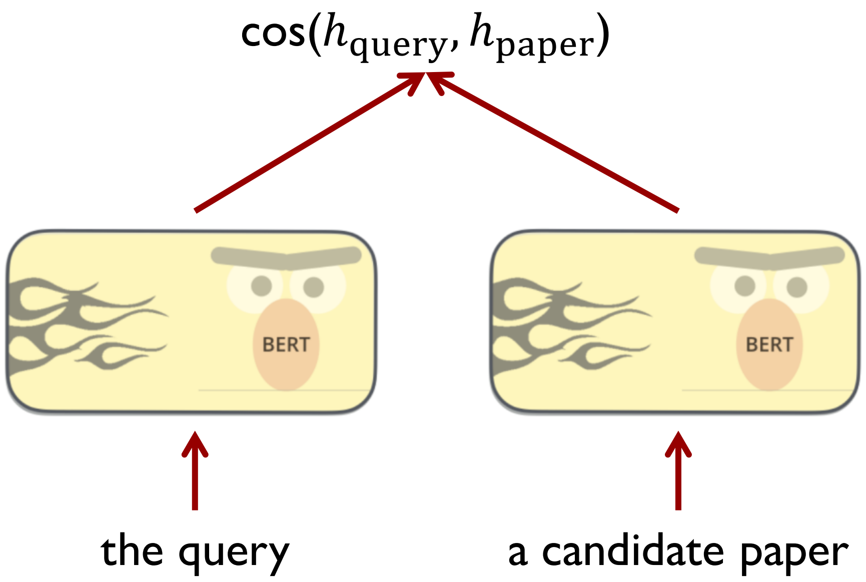
Scientific LLMs: Encoder-Based



Scientific LLMs help literature retrieval.

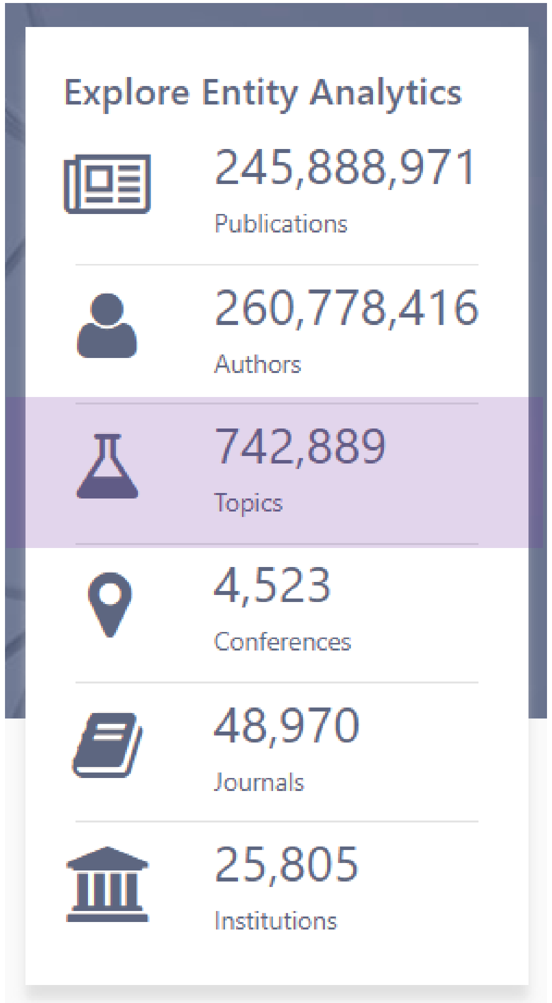
- Given a query (e.g., a scientific claim), how to find related papers from a large corpus?





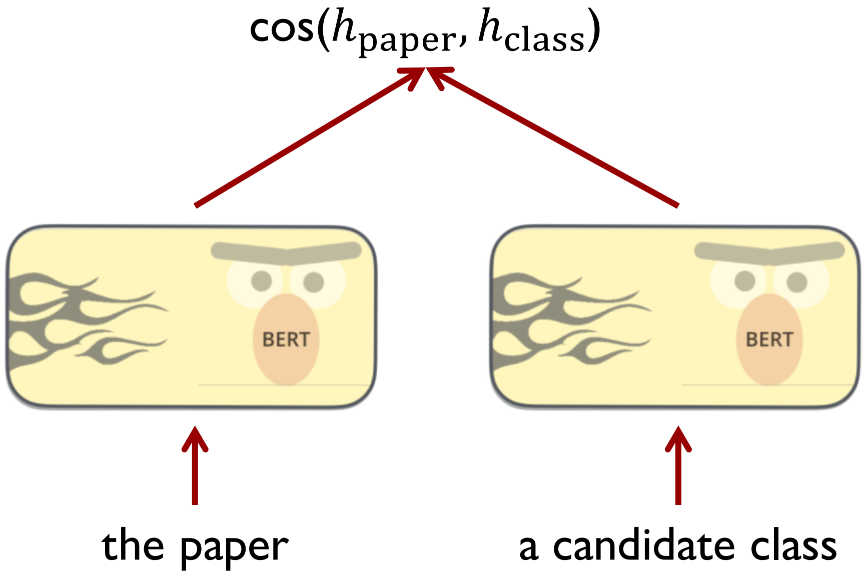
Fact or Fiction: Verifying Scientific Claims. EMNLP 2020.

Scientific LLMs help paper classification.



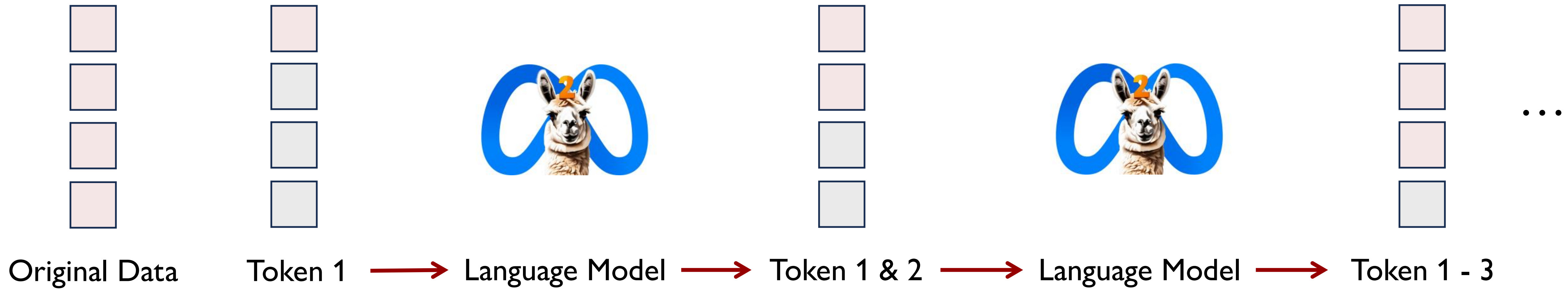
Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study.

Relevant categories: Betacoronavirus, Cardiovascular Diseases, Comorbidity, Coronavirus Infections, Fibrin Fibrinogen Degradation Products, Mortality, Pandemics, Patient Isolation, Pneumonia, …



Metadata-Induced Contrastive Learning for Zero-Shot Multi-Label Text Classification. WWW 2022.

Scientific LLMs: Decoder-Based



- Suppose you have collected a large number of math questions and their answers, …
- What will happen if you train a decoder-based LLM using such data via next token prediction?

Question: A line parallel to $y = 4x + 6$ passes through $(5, 10)$. What is the y-coordinate of the point where this line crosses the y-axis?

Target: Since the line is parallel to $y = 4x + 6$, we know that its slope is 4. Thus the equation of the line is $y = 4x + b$ where b is the y-intercept. Plugging $(5, 10)$ into this gives $10 = 4 \cdot 5 + b \Rightarrow b = \boxed{-10}$ which is what we wanted.

Scientific LLMs help question answering.

- The model will be capable of answering other math questions.

Question: Assume that variance of first n natural numbers is 10 and variance of first m even natural numbers is 16. Compute $m + n$.

Model output: Variance of first n natural numbers is $\frac{n^2-1}{12}$. Variance of first m even natural numbers is $\frac{2^2(m^2-1)}{12}$. So, $\frac{n^2-1}{12} = 10$ and $\frac{2^2(m^2-1)}{12} = 16$. Solving for n and m, we get $n = 11$ and $m = 7$. So, $m + n = 18$.

Question: Prove that for every a, b, $b \neq a$, holds $\frac{a^2+b^2}{2} > \left(\frac{a+b}{2}\right)^2$.

Model output: We have

$$\frac{a^2+b^2}{2} > \left(\frac{a+b}{2}\right)^2$$

$$\frac{a^2+b^2}{2} > \frac{a^2+2ab+b^2}{4}$$

$$2a^2+2b^2 > a^2+2ab+b^2$$

$$a^2-2ab+b^2 > 0$$

$$(a-b)^2 > 0,$$

which is true for all a and b.

Solving Quantitative Reasoning Problems with Language Models. NeurIPS 2022.

Scientific LLMs help knowledge extraction.

Instruction: Extract the product, reactants, reaction type, catalyst, solvent, temperature, and yield from the following paper.

Scientific Paper

… The methyl-substituted porphyrinogens (7e and 7f) were oxidized with chloranil, and meso-unsubstituted porphyrinogens (7g and 7h) were oxidized with 0.1% aqueous $FeCl_3$ in $CHCl_3$ at room temperature to obtain 16π-conjugated systems 5e in 6%, 5f in 7%, 5g in 5%, and 5h in 4% yields. …

Chemical Reaction 1	
Product	5e
Reactants	7e
Reaction Type	oxidation
Catalyst	chloranil
Solvent	$CHCl_3$
Temperature	room
Yield	6%

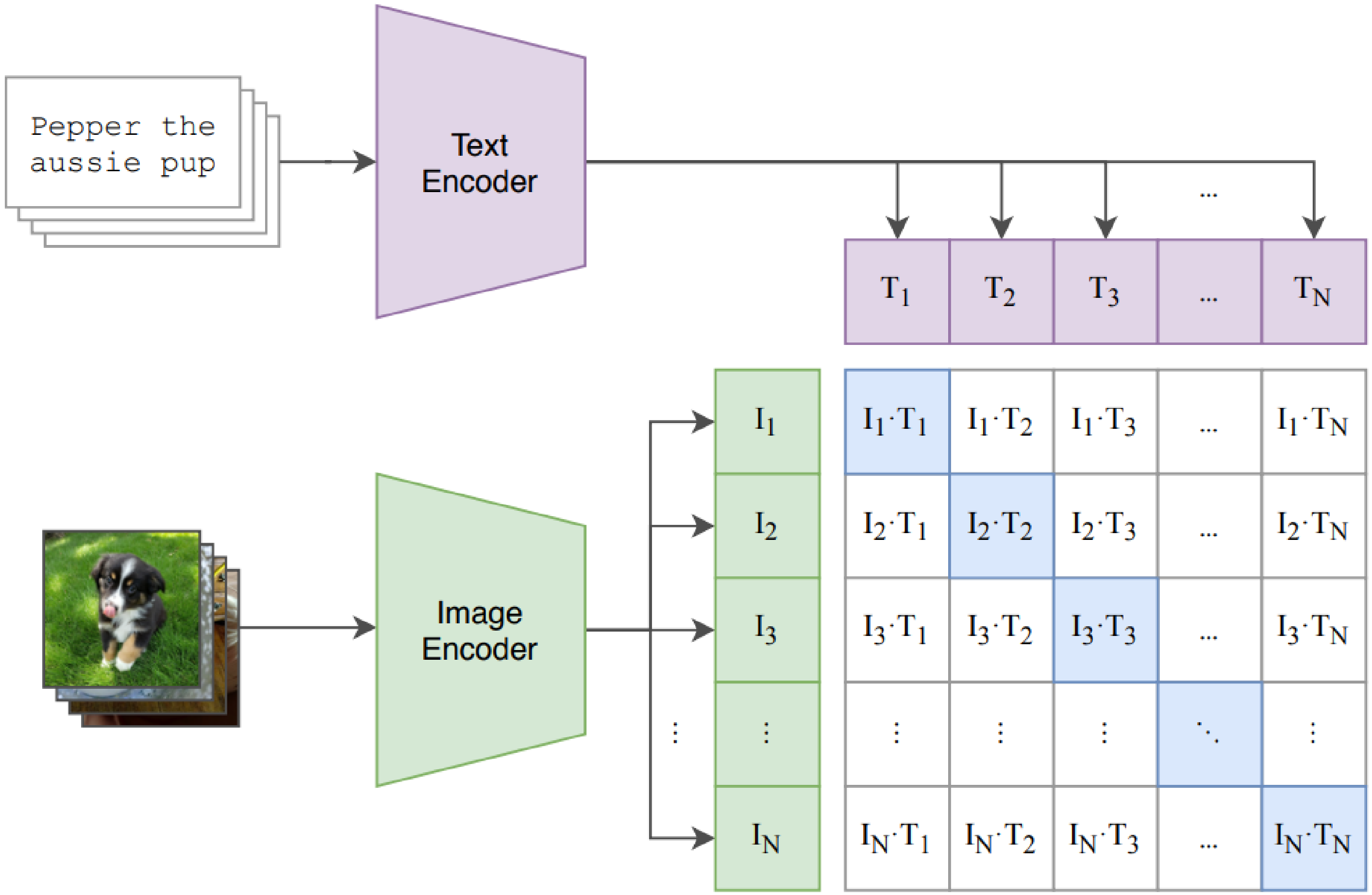
ReactIE: Enhancing Chemical Reaction Extraction with Weak Supervision. ACL 2023.

Overview of Course Contents

- Scientific Large Language Models
  - Encoder-Only, Decoder-Only, Encoder-Decoder
- Fundamental Scientific NLP Tasks
  - Citation Prediction, Literature Retrieval, Question Answering, Knowledge Extraction, Paper Classification
- Scientific Large Vision-Language Models
  - Bioimaging, Geometry, Geography, …
- Scientific Language Models for Other Data Modalities
  - Protein, DNA/RNA, Molecule, Academic Graph, Table, …
- Scientific NLP for Automating Research
  - Idea Generation, Content Generation, Execution, Reviewing, …

Contrastive Language-Image Pre-training (CLIP)

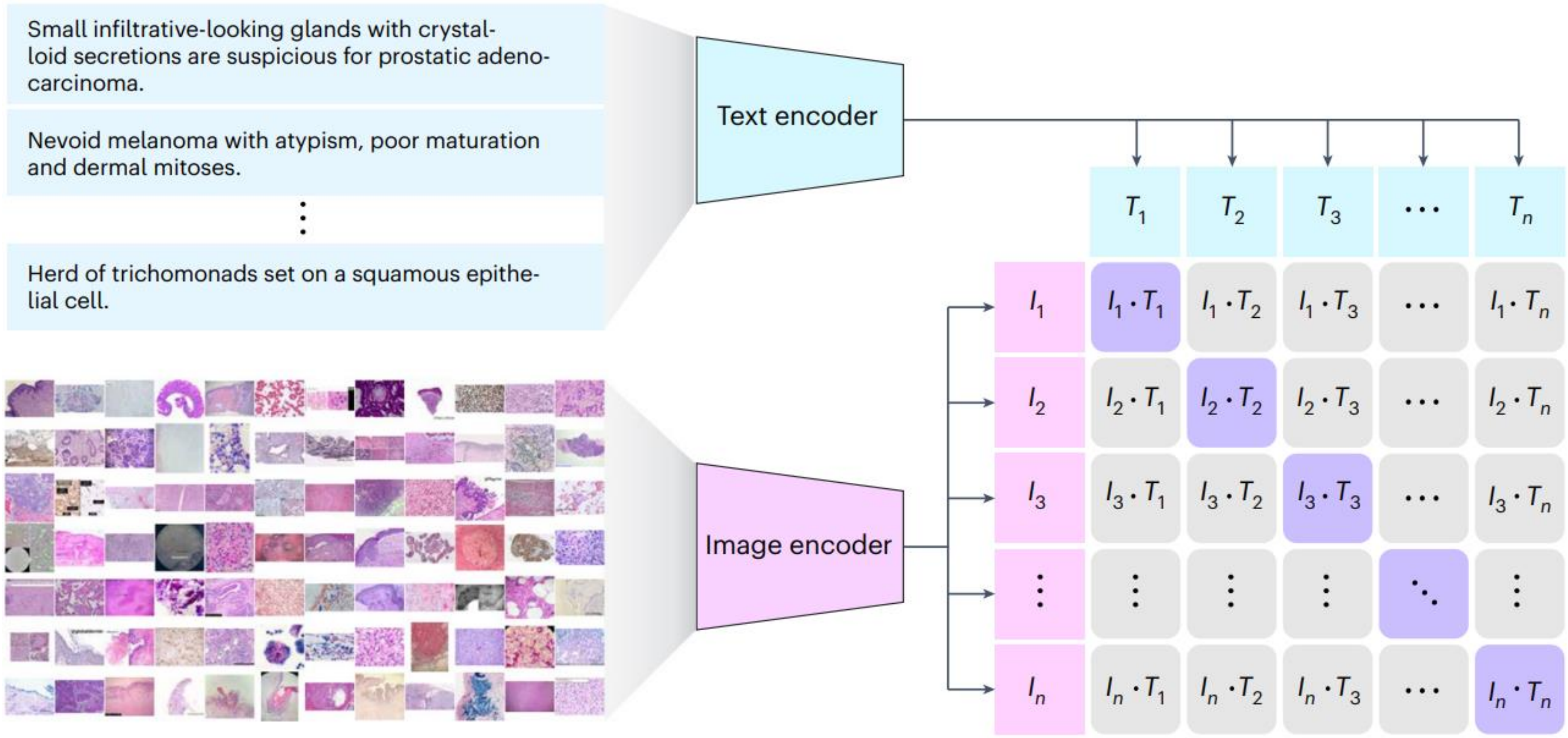
- Suppose you have collected a large number of (image, description) pairs, …
- How to build a model that jointly considers text and images?
- What can CLIP do?
    - Image-to-text retrieval
    - Text-to-image retrieval
    - You can use the text encoder alone to perform text-only tasks.



Learning Transferable Visual Models From Natural Language Supervision. ICML 2021.

Scientific VLMs help image-based diagnosis.

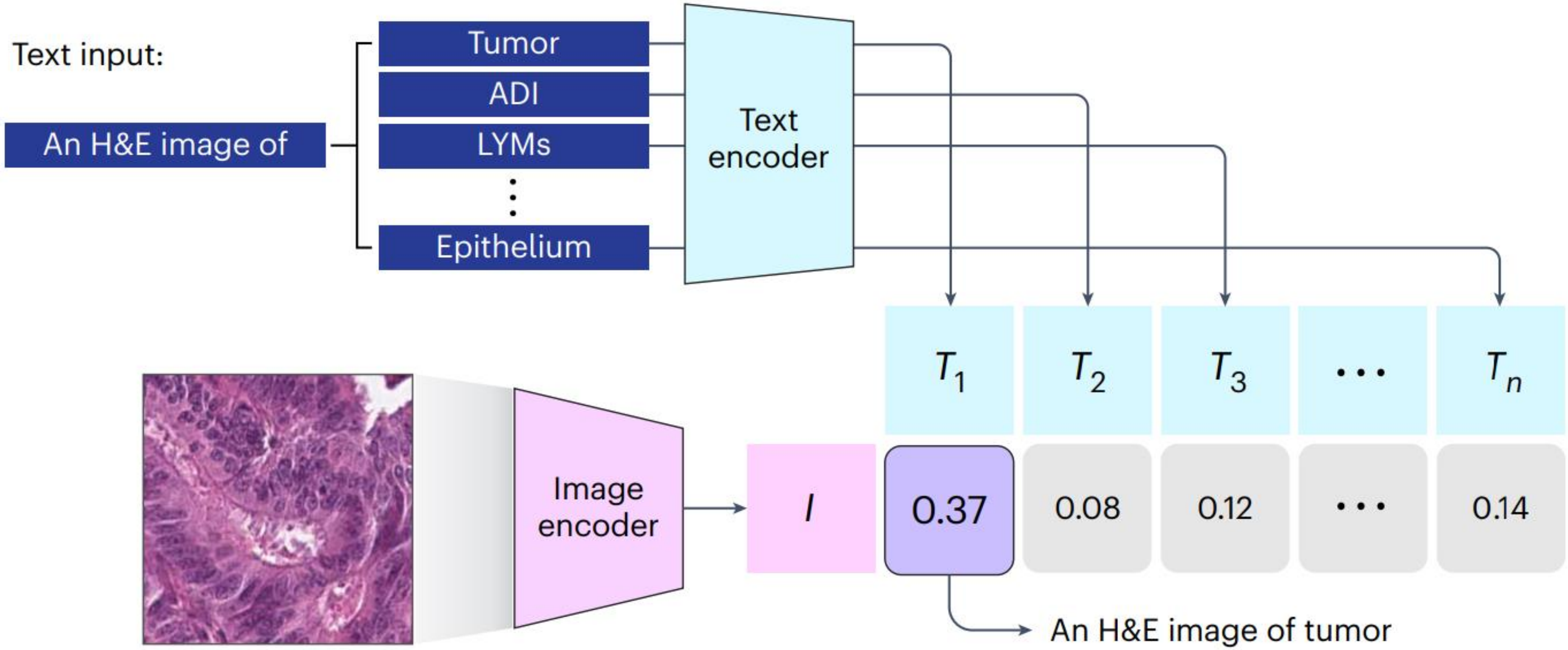
- Train a CLIP model using biomedical images and their associated text



A Visual–Language Foundation Model for Pathology Image Analysis using Medical Twitter. Nature Medicine 2023.

Scientific VLMs help image-based diagnosis.

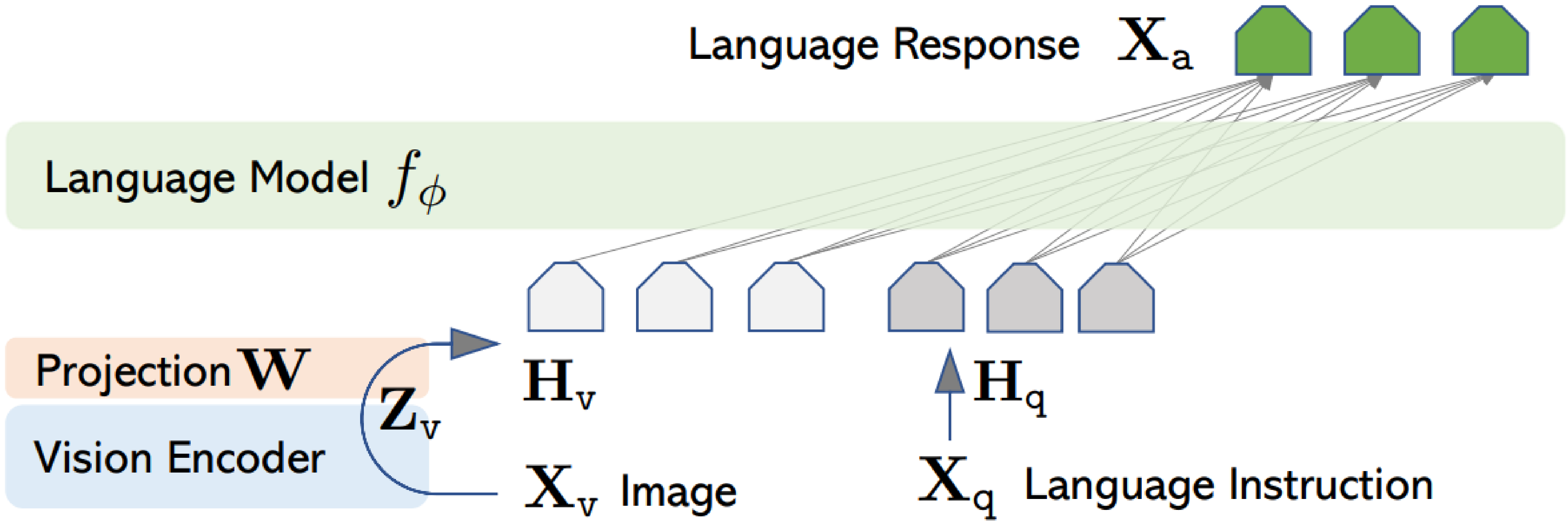
- Perform image-based diagnosis by enumerating all possibilities with a text template.



A Visual–Language Foundation Model for Pathology Image Analysis using Medical Twitter. Nature Medicine 2023.

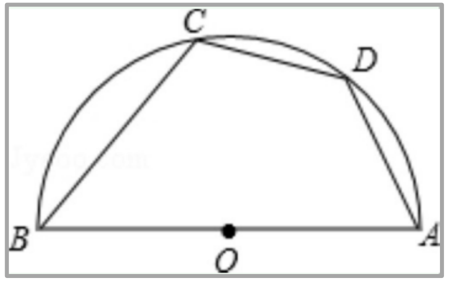
Adding Images into a Decoder-Based Architecture – LLaVA

- Project images onto several vision tokens.
- Prepend vision tokens to text tokens for next token prediction.



Visual Instruction Tuning. NeurIPS 2023.

Scientific VLMs help geometric problem solving.



As shown in the figure, AB is the diameter of the semicircle, angle ABC = 50.0, point D is the midpoint of arc AC, then angle DAB is ( ) degrees.

G-LLaVA

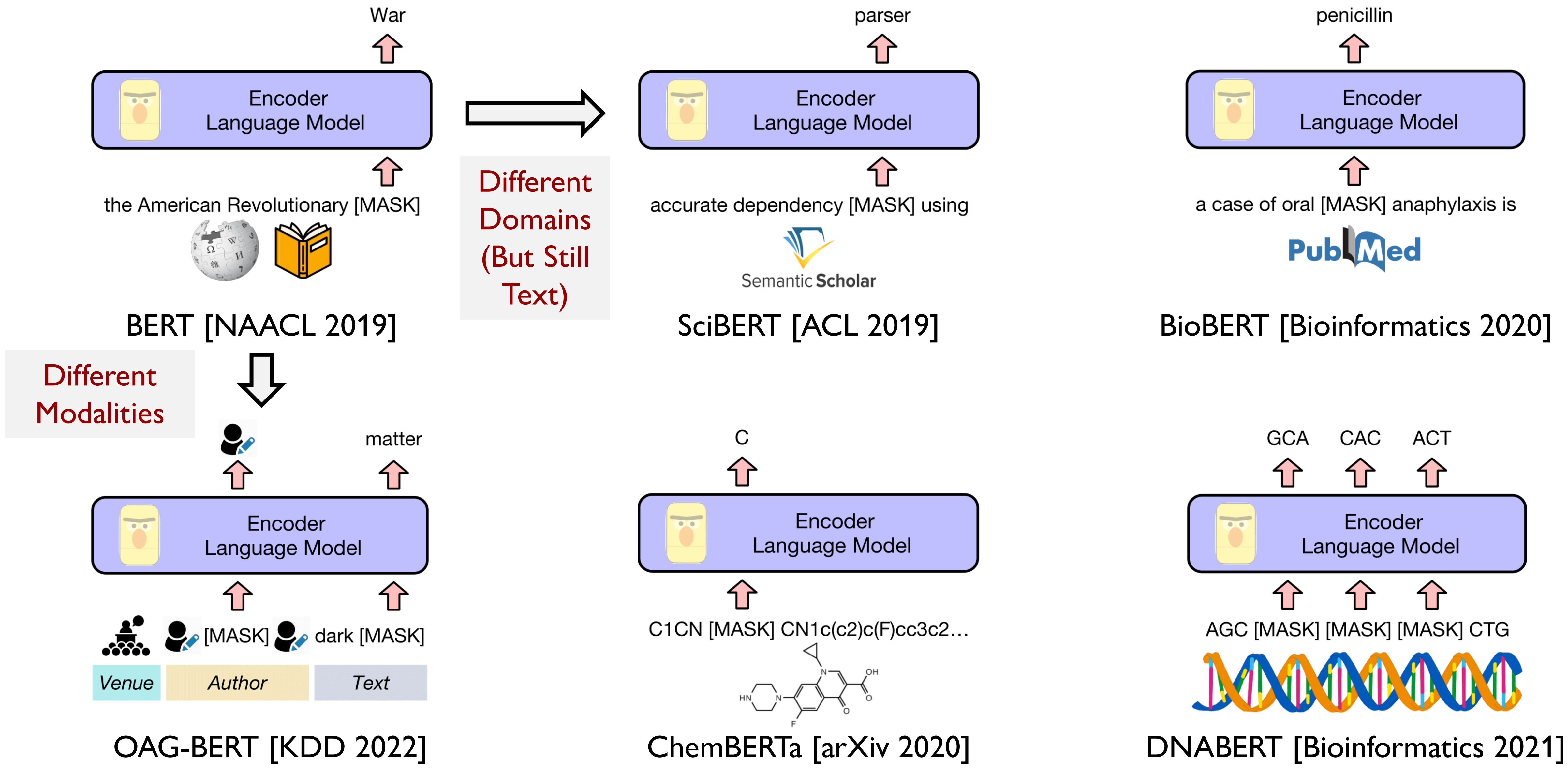
Connecting BD, as shown in the diagram. Because point D is the midpoint of segment AC, that is, CD equals AD. Therefore, angle ABD equals angle CBD. And angle ABC is 50 degrees, so angle ABD is half of 50 degrees, which is 25 degrees. Since AB is the diameter of the semicircle, angle ADB is 90 degrees. Therefore, angle DAB is 90 degrees minus 25 degrees, which is 65 degrees.

G-LLaVA : Solving Geometric Problem with Multi-Modal Large Language Model. arXiv 2023.

Overview of Course Contents

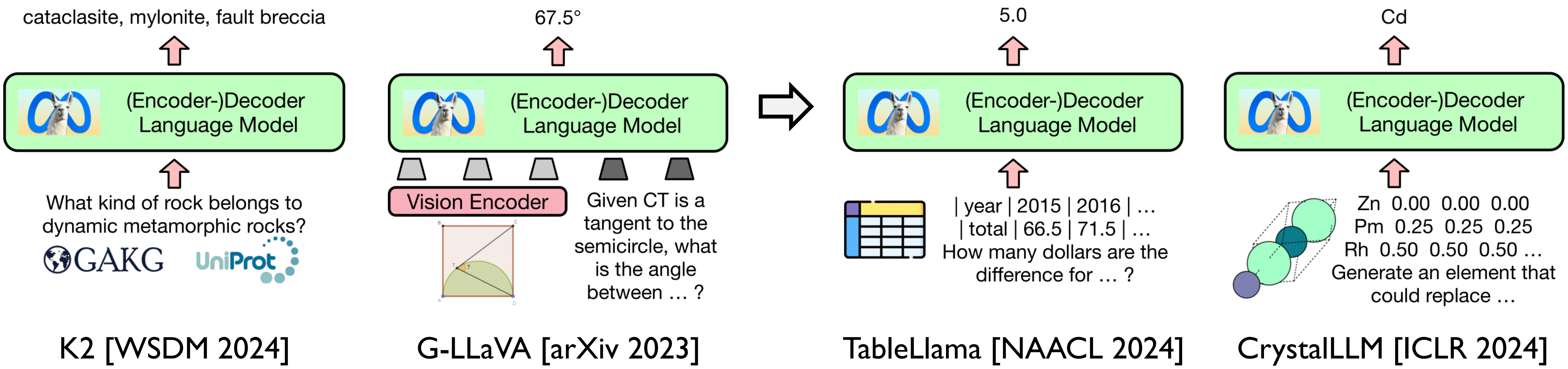
- Scientific Large Language Models
  - Encoder-Only, Decoder-Only, Encoder-Decoder
- Fundamental Scientific NLP Tasks
  - Citation Prediction, Literature Retrieval, Question Answering, Knowledge Extraction, Paper Classification
- Scientific Large Vision-Language Models
  - Bioimaging, Geometry, Geography, …
- Scientific Language Models for Other Data Modalities
  - Protein, DNA/RNA, Molecule, Academic Graph, Table, …
- Scientific NLP for Automating Research
  - Idea Generation, Content Generation, Execution, Reviewing, …

Extending Encoder Architecture to Other Modalities



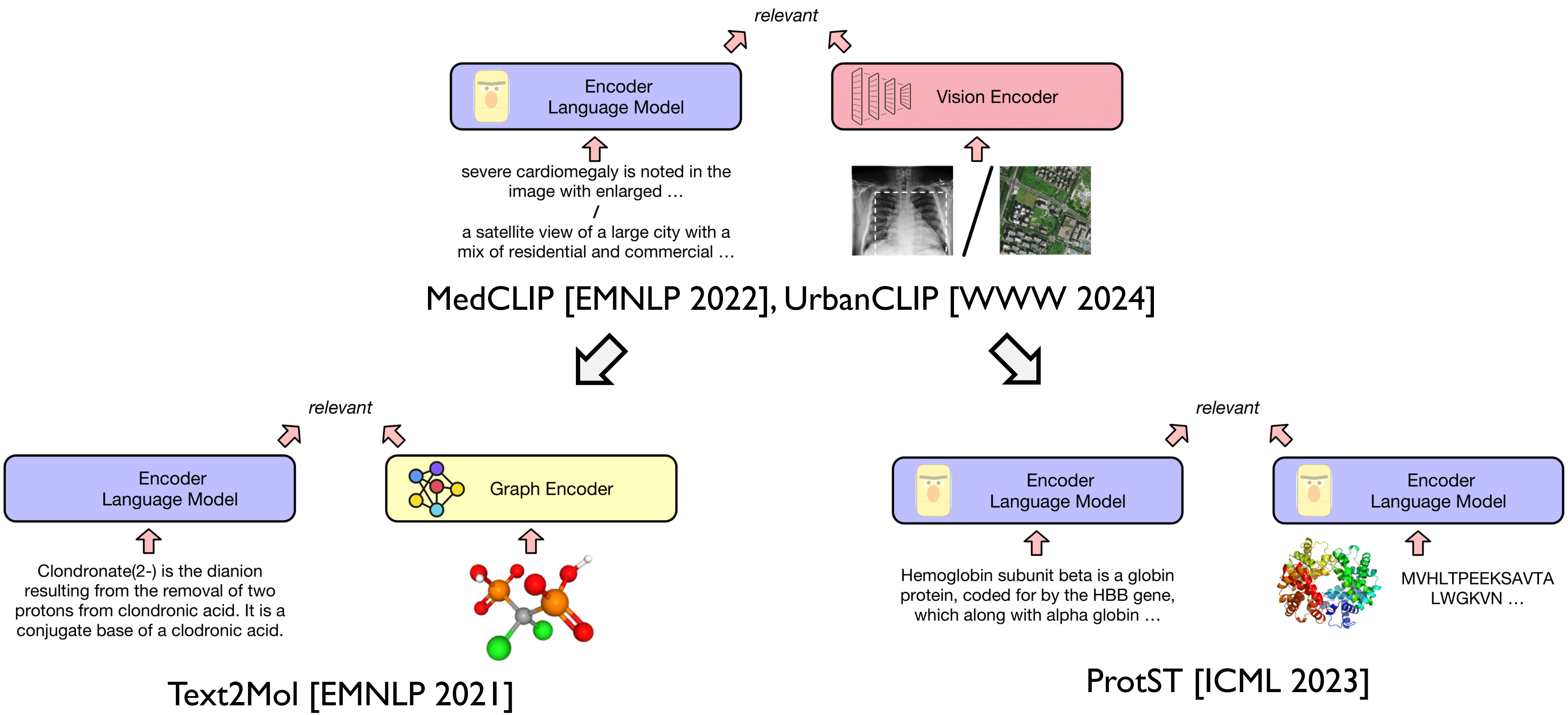
Extending Decoder Architecture to Other Modalities

- Linearizing scientific data → Next token prediction (→ Instruction Tuning → …)

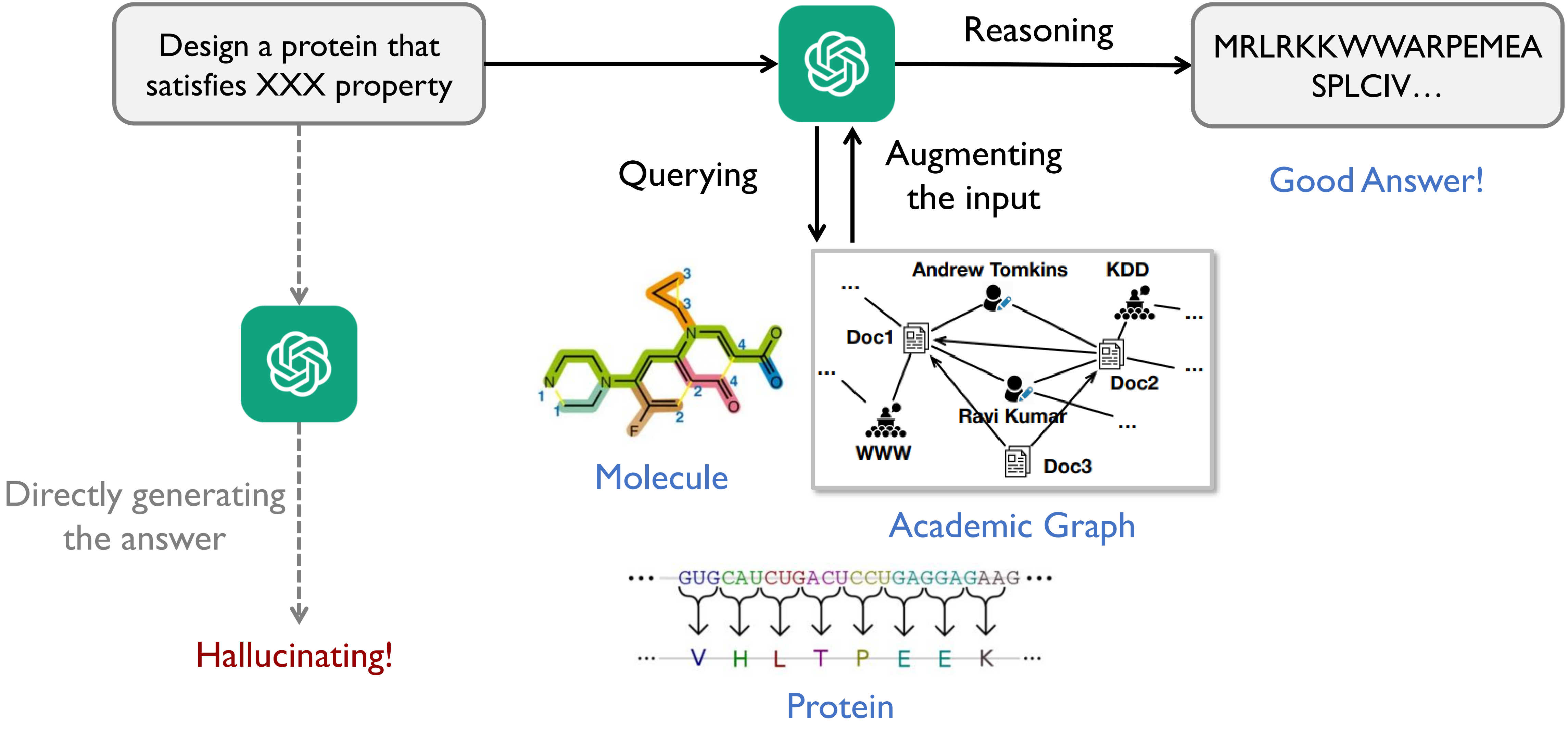


Extending CLIP Architecture to Other Modalities

- Mapping text and relevant scientific data closer in the latent space



Other modalities enhance the trustworthiness of scientific LLMs.



Taking Academic Graph as an Example

You

What is the most cited paper in WWW 2017? Please list its title and authors.

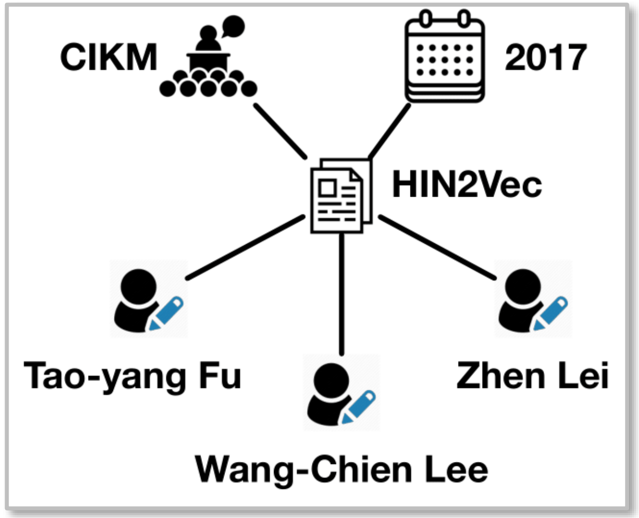
ChatGPT

As of my last update in January 2022, I don't have the real-time citation information, but one of the highly cited papers from WWW 2017 is:

Title: "HIN2Vec: Explore Meta-paths in Heterogeneous Information Networks for Representation Learning"

Authors: Cheng Yang, Zhiyuan Liu, Deli Zhao, Maosong Sun, Edward Y. Chang

Please verify the latest citation counts and updates from reliable sources or academic databases for the most current information.



HIN2Vec was published in CIKM 2017.

HIN2Vec was written by Tao-yang Fu, Wang-Chien Lee, and Zhen Lei.

ChatGPT 3.5, queried on January 23, 2024

Taking Academic Graph as an Example

What is the most cited paper in WWW 2017? Please list its title and authors.

As of January 2025, the most cited paper from the 26th International World Wide Web Conference (WWW 2017) is:

Title: "Attention Is All You Need"

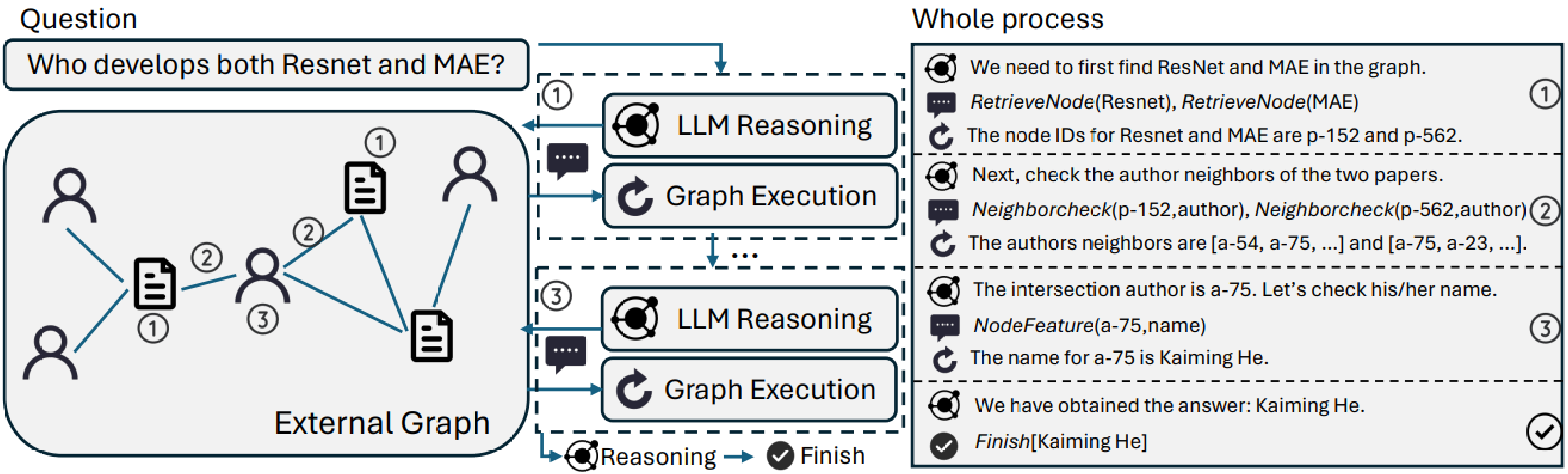
Authors: Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin

This seminal paper introduced the Transformer architecture, which has become foundational in the field of natural language processing and has significantly influenced subsequent research and applications.

ChatGPT 4o, queried on January 7, 2025

Transformer was published in NeurIPS 2017.

LLMs with Academic Graphs



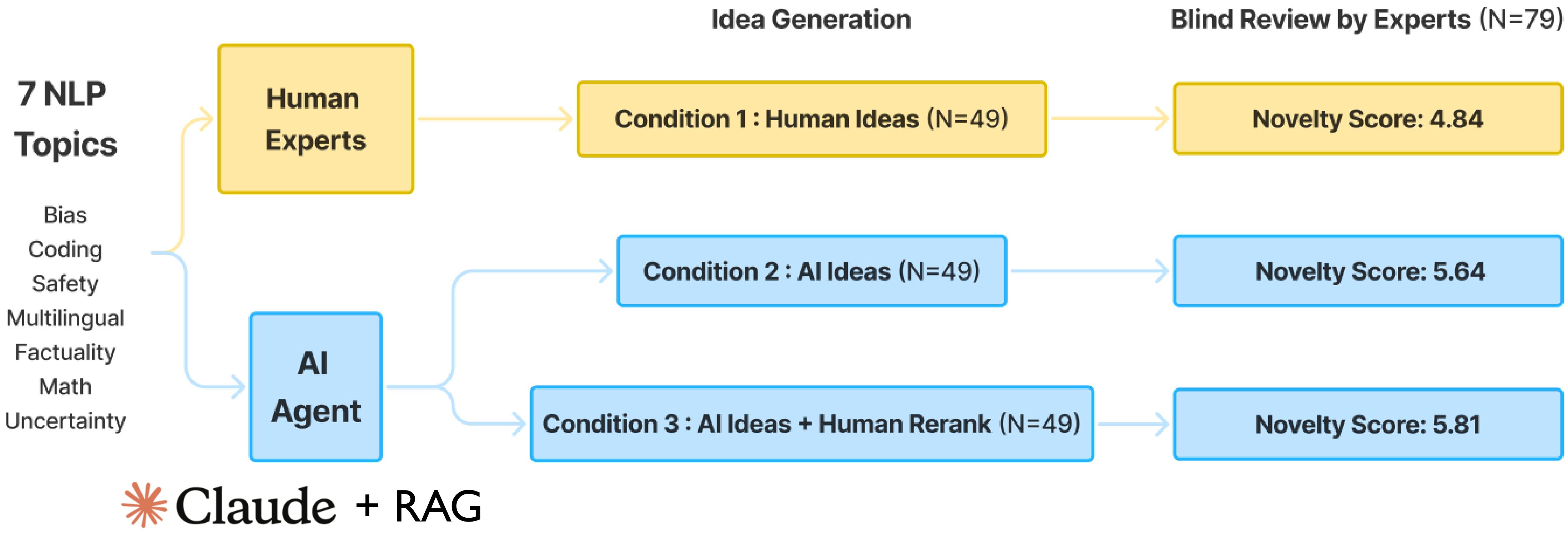
Graph Chain-of-Thought: Augmenting Large Language Models by Reasoning on Graphs. ACL 2024 Findings.

Overview of Course Contents

- Scientific Large Language Models
  - Encoder-Only, Decoder-Only, Encoder-Decoder
- Fundamental Scientific NLP Tasks
  - Citation Prediction, Literature Retrieval, Question Answering, Knowledge Extraction, Paper Classification
- Scientific Large Vision-Language Models
  - Bioimaging, Geometry, Geography, …
- Scientific Language Models for Other Data Modalities
  - Protein, DNA/RNA, Molecule, Academic Graph, Table, …
- Scientific NLP for Automating Research
  - Idea Generation, Content Generation, Execution, Reviewing, …

LLMs for Idea Generation

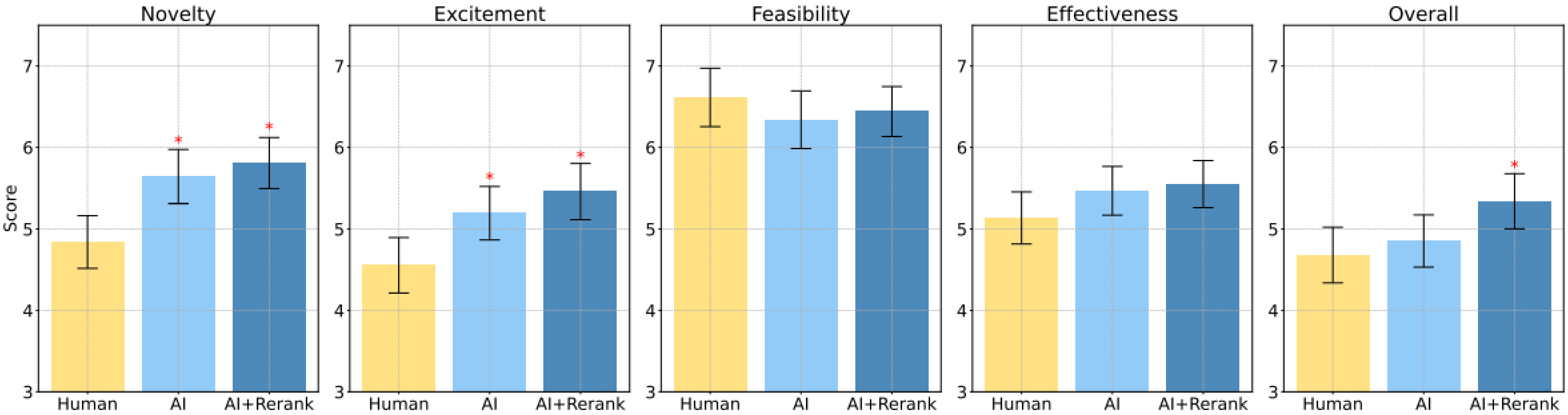
- AI ideas are judged as significantly more novel than human ideas.



Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers. arXiv 2024.

LLMs for Idea Generation

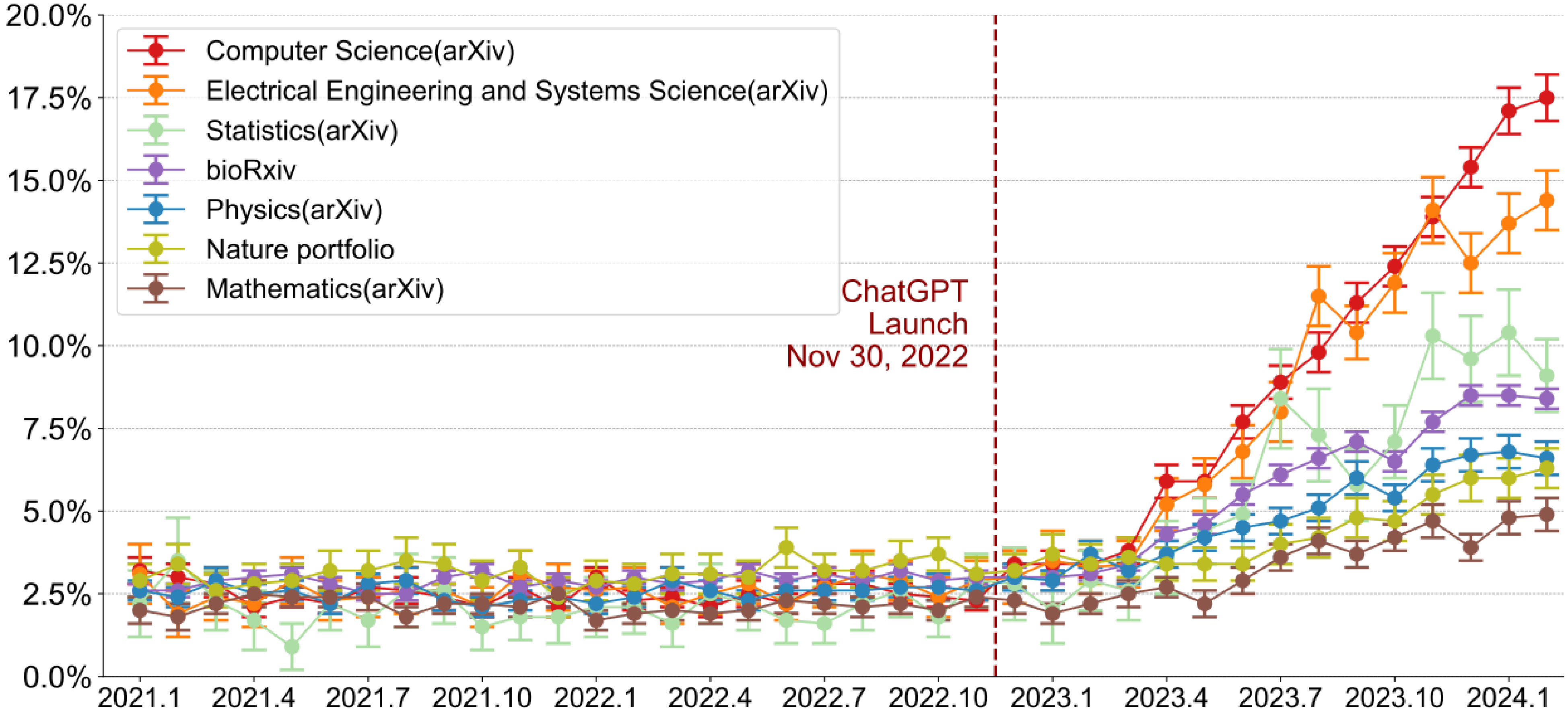
- Human ideas are judged as more feasible than AI ideas.



Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers. arXiv 2024.

LLMs for Paper Writing

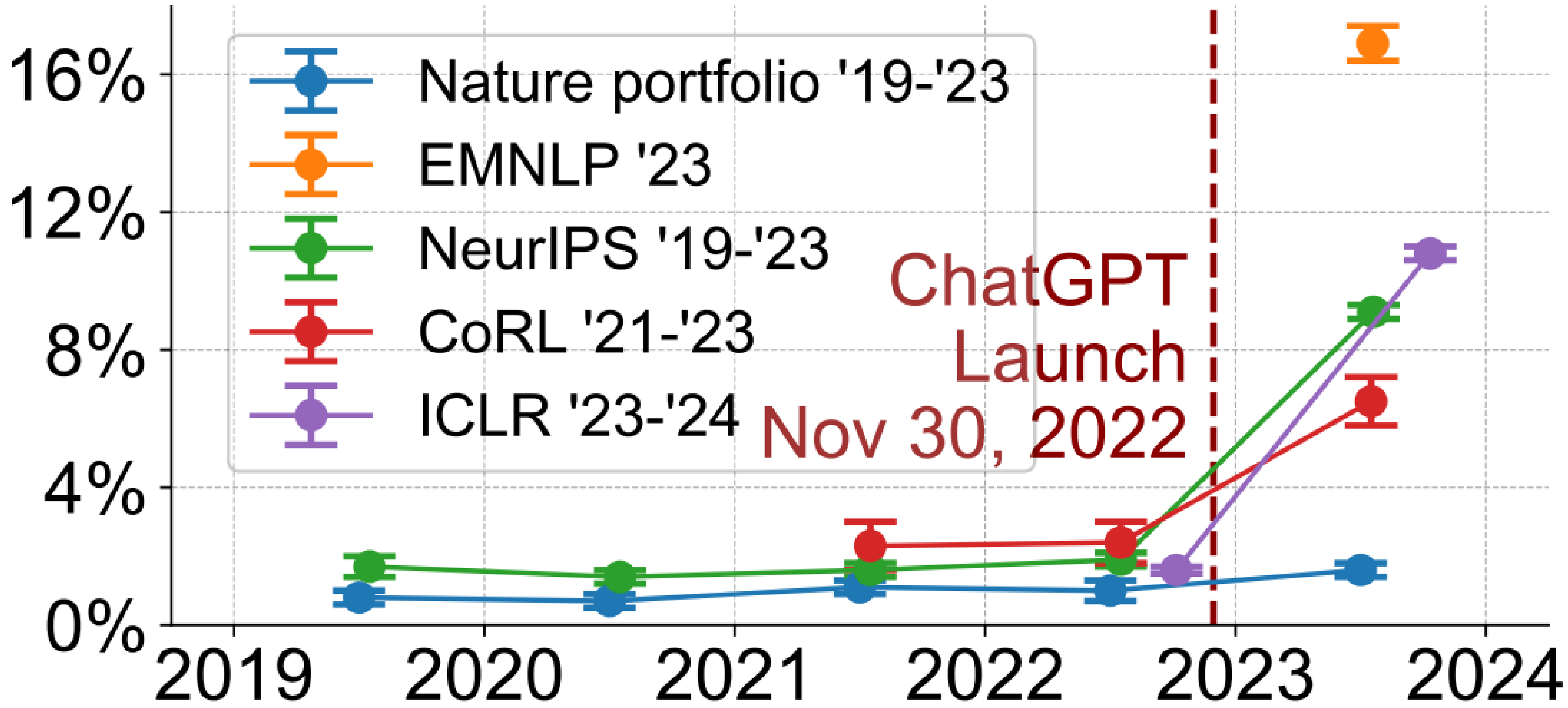
- Estimated fraction of LLM-modified sentences in academic papers over time



Mapping the Increasing Use of LLMs in Scientific Papers. COLM 2024.

LLMs for Reviewing

- Estimated fraction of LLM-modified sentences in paper reviews over time



Mapping the Increasing Use of LLMs in Scientific Papers. COLM 2024.

LLMs for Building Autonomous Agents

- LLMs use tools to search the Web, plan the synthesis, and execute the synthesis.
- Output of the previous tool → LLM → Input of the next tool

Task: Find and synthesize a thiourea organocatalyst which accelerates a Diels-Alder reaction.

First, I need to find a thiourea organocatalyst that can accelerate the Diels-Alder reaction. I will perform a web search to find a suitable catalyst.
Web Search tool: Schreiner's thiourea catalyst

Now, I will obtain the SMILES. **Name2Smiles tool:**
FC(F)(F)c1cc(NC(=S)Nc2cc(C(F)(F)F)cc(C(F)(F)F)c2)cc(C(F)(F)F)c1

I will plan a synthesis for Schreiner's thiourea catalyst.
SynthesisPlanner tool: detailed synthesis plan

I will execute the synthesis.
SynthesisExecuter tool: successful synthesis.

Connection with physical world →

c RoboRXN synthesis platform

Augmenting Large Language Models with Chemistry Tools. Nature Machine Intelligence 2024.

Thank You!

Course Website: https://yuzhang-teaching.github.io/CSCE689-S25.html