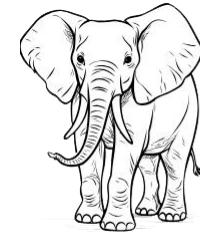




清华大学
Tsinghua University



ACL 2024 Sharing

Zhu Liu

2024.09.26

ACL 2024

- 会议时间：8月11日到16日
(17,18 还有2天workshop)
- 会议地点：泰国曼谷的
Centara Grand and Bangkok
Convention Centre at
CentralWorld





曼谷整体印象

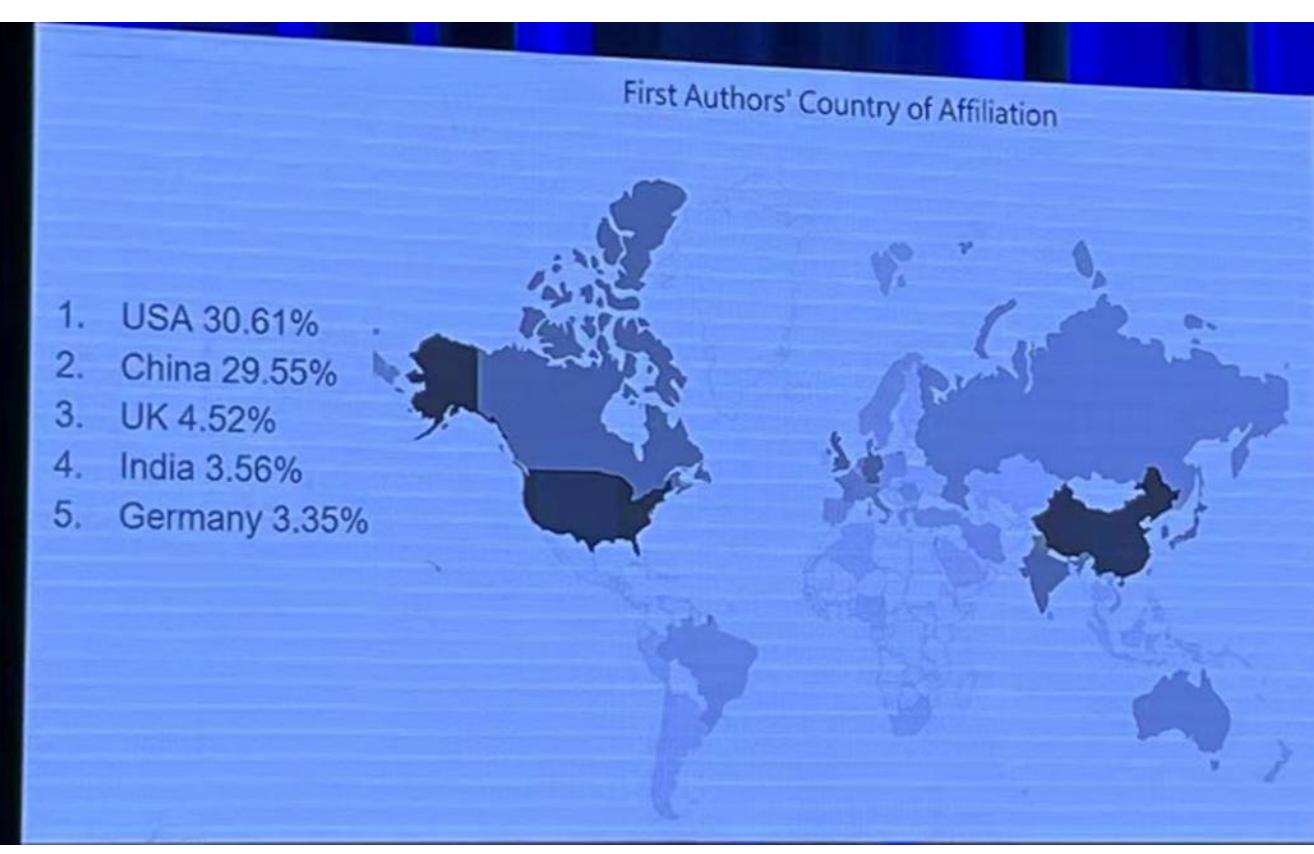
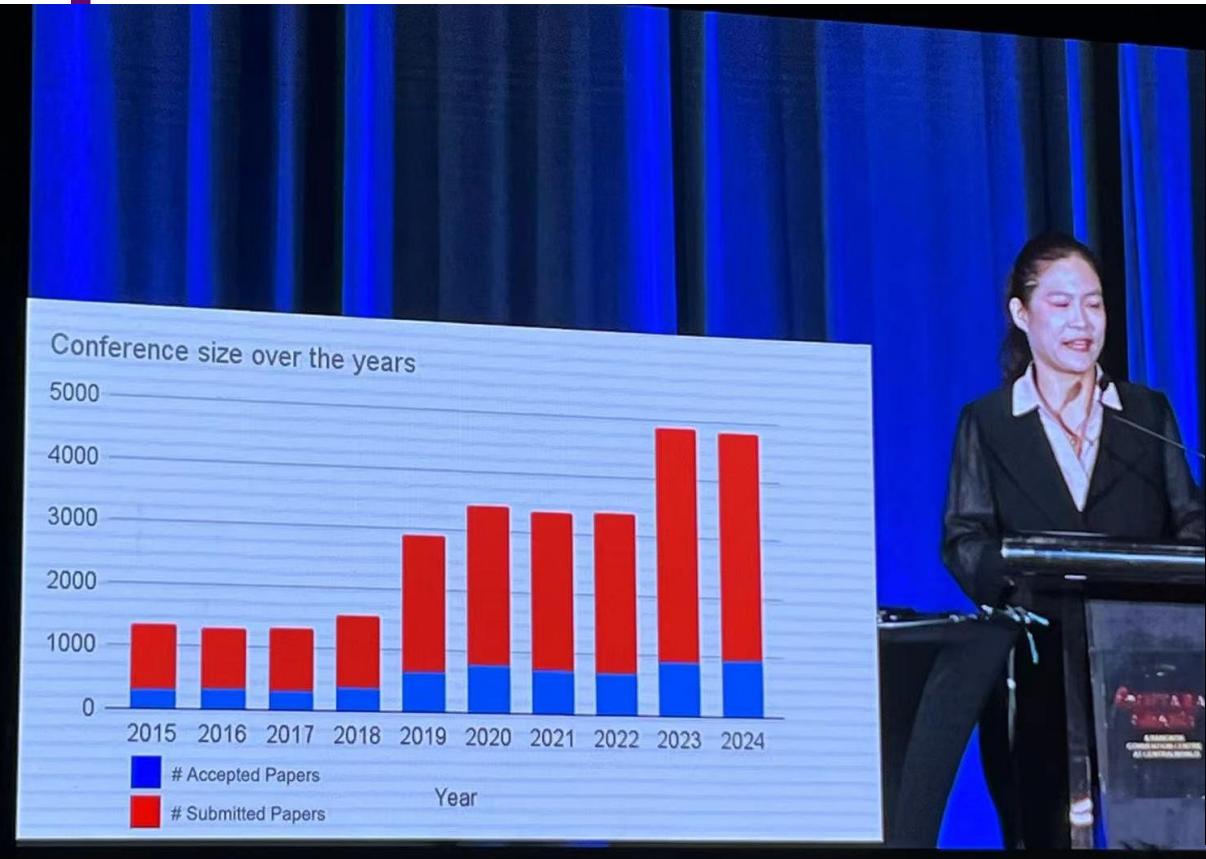
- 传统和现代的交融
 - 传统：很多佛教元素、保留王室、神秘的建筑
 - 现代：高楼大厦、便捷的购物体验...
- 天气湿热
- 饮食偏辛辣
- 一些可能吸引办会的点：国际化程度高、交通便捷、签证方便、英语普及度较高、很安全

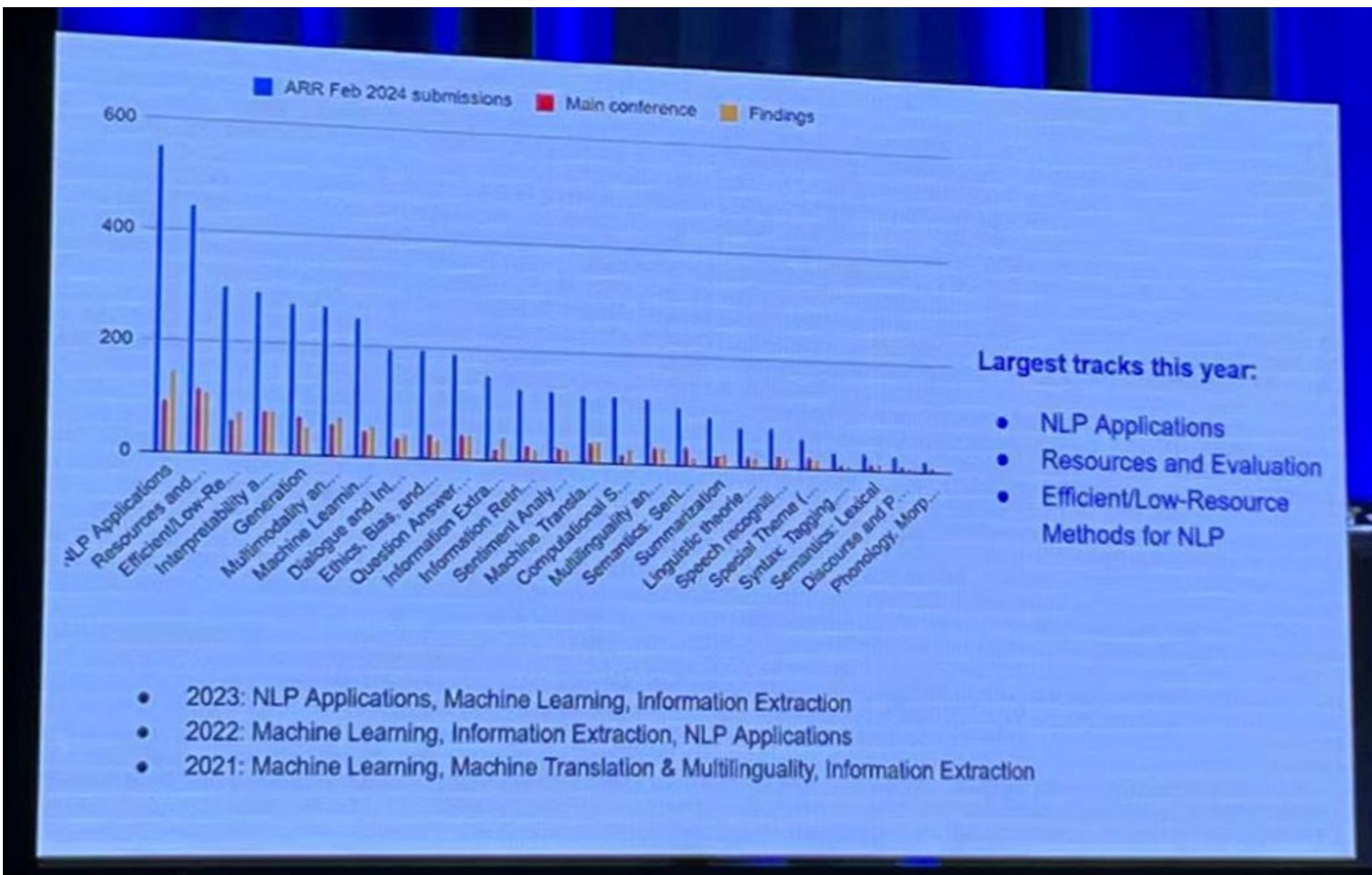
议程

日期	时间	活动内容
8月10日	14:00-19:00	注册
8月11日	07:30-20:30	注册、教程、欢迎招待会
8月12日	07:30-17:30	注册、主题演讲、口头报告/海报/演示
8月13日	08:30-16:30	注册、主题演讲、圆桌讨论、社交晚宴
8月14日	08:30-16:30	注册、主题演讲、奖项与闭幕
8月15-16日	08:00-16:30	注册、研讨会

ACL 2024 介绍

- 主会
 - 约5000篇投稿， 72个SAC， 716个AC， 4208个审稿人
 - 940篇main (21.3%) , 975篇findings (22.1%) , 6 JCL, 31 TACL
 - 3个主旨演讲， 1个圆桌讨论
- 其他
 - 18个workshop, 6个教程, 38个demos和60篇SRW的论文





What's New This Year



- All papers presented as posters, some papers also as orals
- All findings papers can present (as posters)
- Virtual day **after** the conference (next week: Thursday, August 22)
- Non-publicized paper award
- Theme: Open science, open data, and open models for reproducible NLP

主旨演讲

- Does In-Context-Learning Offer the Best Tradeoff in Accuracy, Robustness, and Efficiency for Model Adaptation?
- Can LLMs Reason and Plan?
- Are LLMs Narrowing Our Horizon? Let's Embrace Variation in NLP!

主旨演讲

- Does In-Context-Learning Offer the Best Tradeoff in Accuracy, Robustness, and Efficiency for Model Adaptation?
- 主讲人：Sunita Sarawagi
 - Professor, IIT Bombay
 - sequence models for text and time-series, domain adaptation, effective human intervention in learning, graphical models and structured learning



Background

- The model adaptation problem: from one domain (training set) to another (test set)
- Challenges:
 - Accuracy (overfit vs. no change)
 - Robustness
 - Efficiency
- Related topics: domain adaptation/transfer learning/few-shot learning
- Methods: Fine-tuning and its variants; Mixture of experts; Task-vectors; Matching-based methods

Evaluation

(small T)	Fine-tuning	MoE	Task vectors	Matching
Accuracy	✓✓	✓✓✓	✓	✓
Robustness	✓	✓✓	✓✓✓	✓✓
Efficiency		✓	Online ✓✓✓	Online ✓✓✓
Adapt	✓			
Test	✓✓✓	✓✓✓	✓✓✓	✓
Others		Difficult to train		

LLM-era

- In-context learning (ICL): Adaptation is just a forward pass
- Why does it work?
 - H1: Transformers implement gradient descent algorithm over IC examples
 - H2: IC examples recognize tasks in pre-training, e.g. via **task vectors**
 - H3: Self-attention implements **matching-based** adaptation via induction heads

Re-evaluation

(small T)	Fine-tuning	ICL
Accuracy	✓✓	✓✓
Robustness	✓	✓
Efficiency		Online
Adapt	✓	✓✓✓
Test	✓✓✓	✓✓
Others		Ease of use in multi-tenancy models

Improvement:

- Structuring of prompts
- Pre-training strategies
- Fine-tuning
- Example retrieval
- Model architectures

Can LLMs Really Reason and Plan?

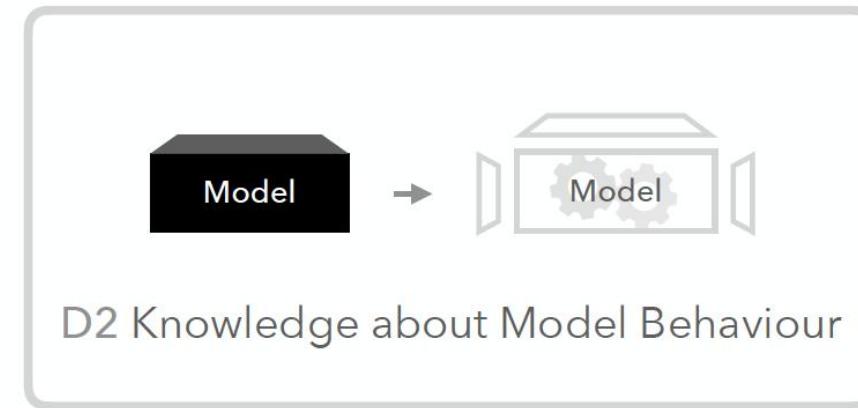
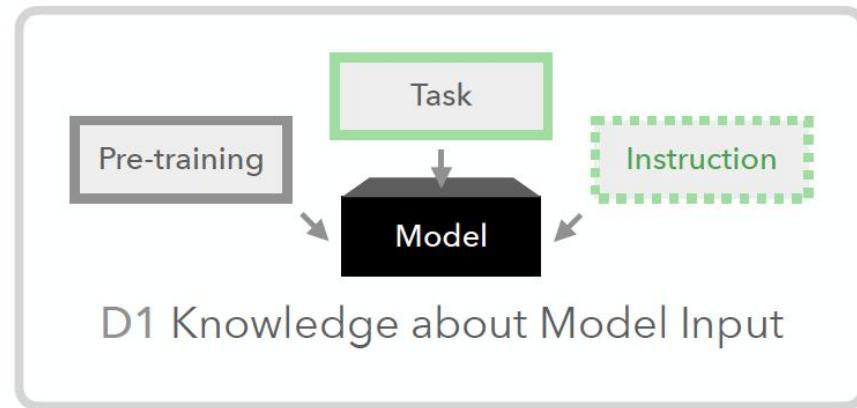
- Speaker: Subbarao Kambhampati
- Arizona State University
- 关注的问题：LLMs在推理过程中会出现幻觉等问题
- 作者对LLM常用的微调和提示语进行了很多测试
- 结论：LLMs的训练和使用方式并没有显示出它们能够进行原则性推理的迹象。



Are LLMs Narrowing Our Horizon? Let's Embrace Variation in NLP!

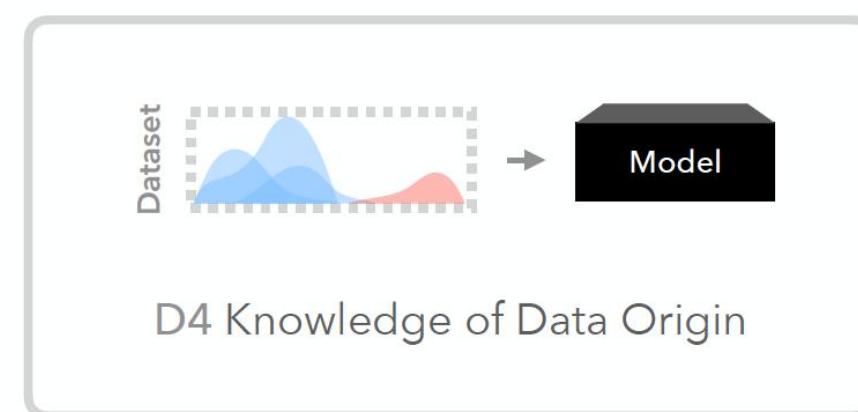
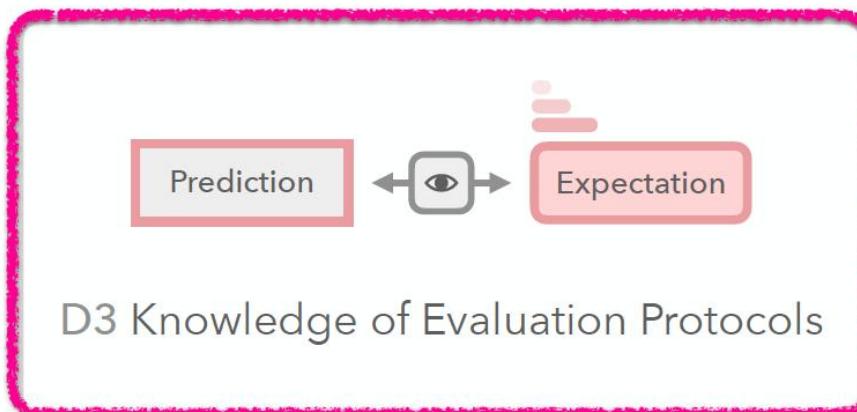
- Speaker: Barbara Plank LMU Munich & IT University of Copenhagen
- LLMs & Trust: A Short Look at AI history
- Regain trust from four aspects:





Trust arises from knowledge of origin as well as from knowledge of functional capacity.

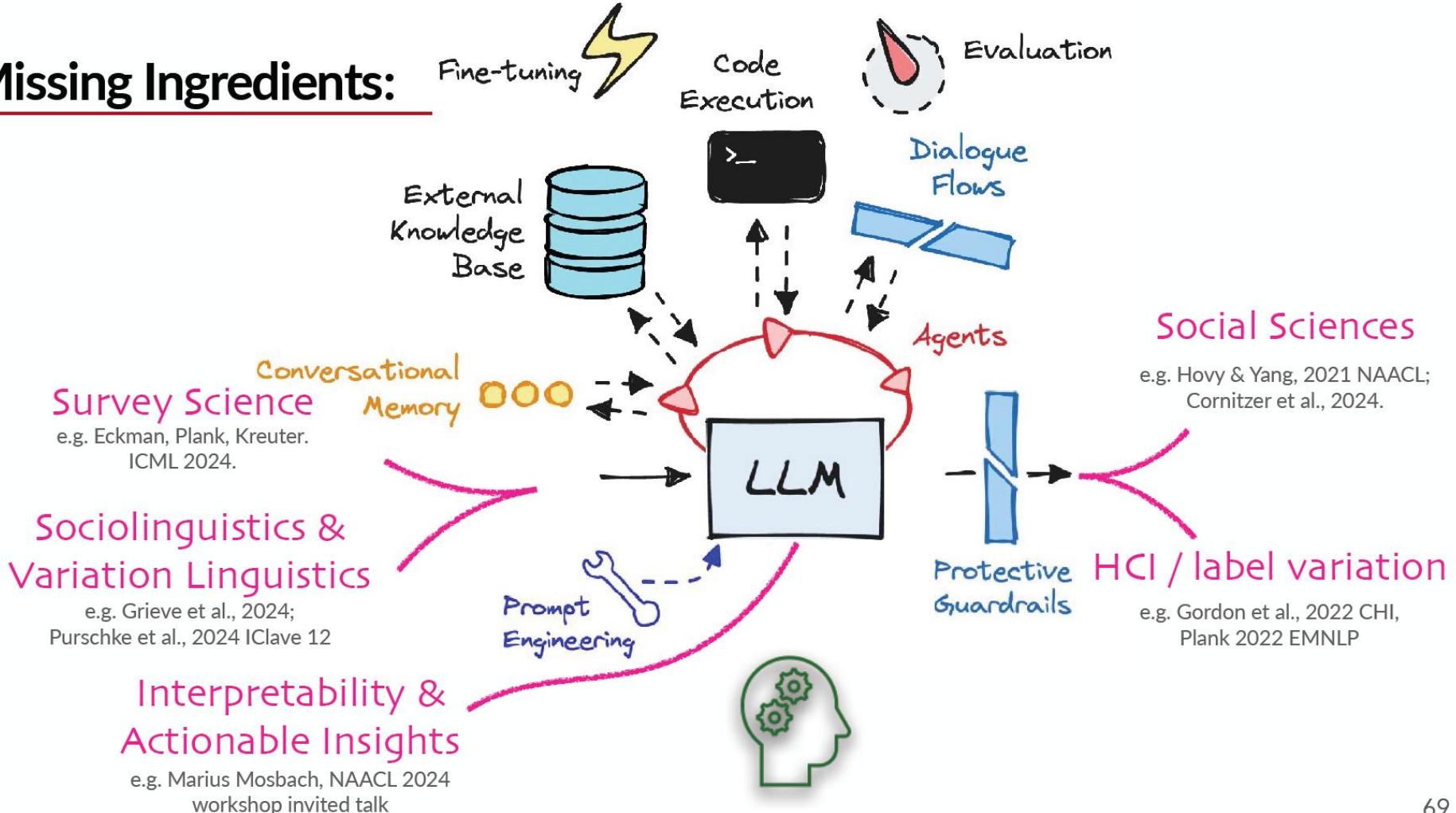
Trustworthiness - Working Definition by David G. Hays, 1979



We need to embrace variation holistically

- Inputs: linguistic variation, low-resource languages & dialects
- Outputs: human label variation as signal (not error) [Uncertainty]
- Research: Language as the Bridge

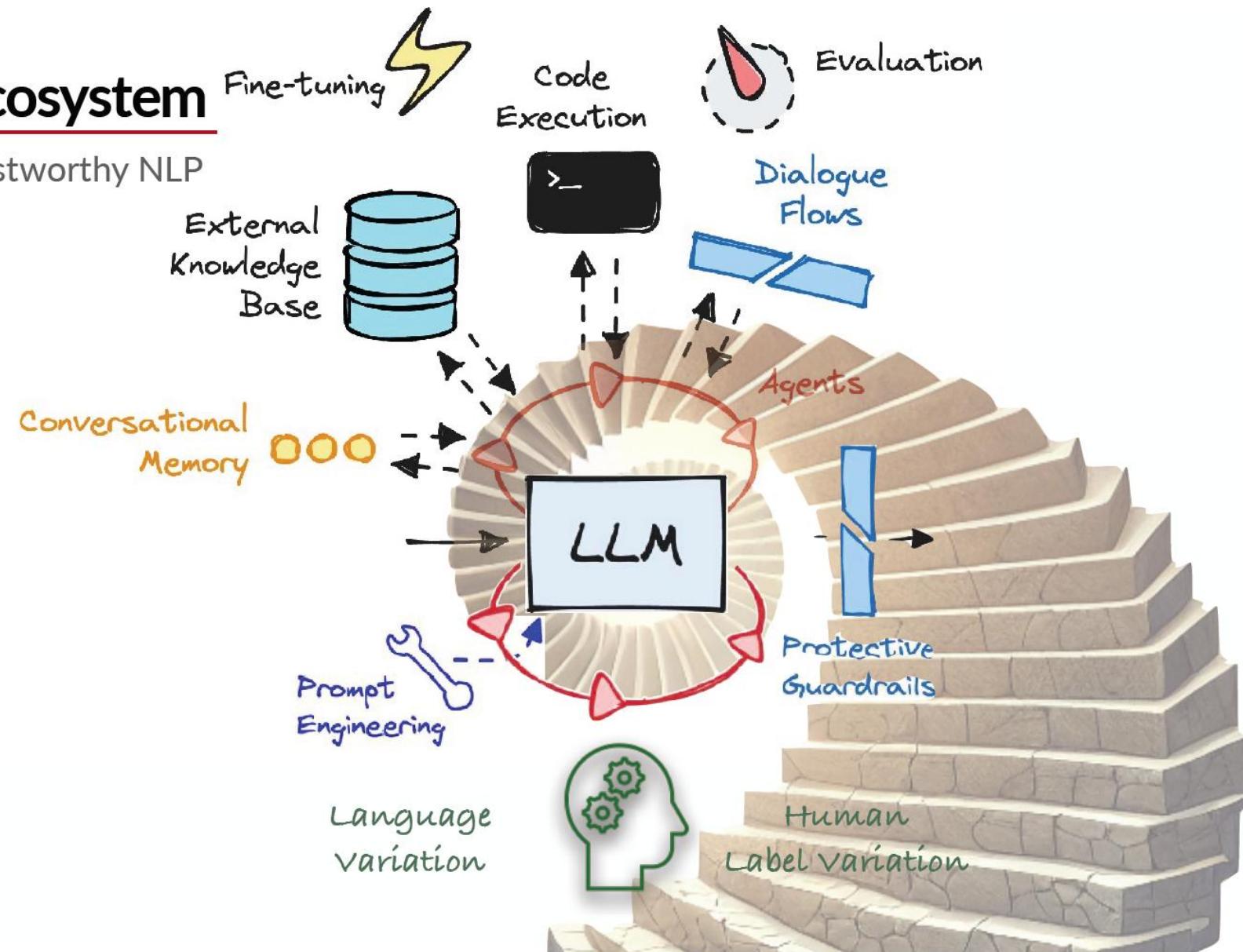
Missing Ingredients:



69

Trust LLM Ecosystem

Human-facing, Trustworthy NLP



Tutorials

- **Tutorial 1– Computational Linguistics for Brain Encoding and Decoding: Principles, Practices and Beyond**

Room : Lotus 1-4 (Level 22)

Jingyuan Sun, and Shaonan Wang, and Zijiao Chen, and Jixing Li, and Marie-Francine Moens

- **Tutorial 2 - Automatic and Human-AI Interactive Text Generation (with a focus on Text Simplification and Revision)**

Room : Lotus 5-7 (Level 22)

Yao Dou, and Philippe Laban, and Claire Gardent, and Wei Xu

- **Tutorial 3 - Vulnerabilities of Large Language Models to Adversarial Attacks**

Room : World Ballroom B (Level 23)

Yu Fu, Erfan Shayegan, and Md. Mamun Al Abdullah, and Pedram Zaree, and Nael Abu-Ghazaleh, and Yue Dong

Tutorials

14:00 - 17:30

Tutorial 4 - 6

- **Tutorial 4 – Computational Expressivity of Neural Language Models**

Room: Lotus 1-4 (Level 22)

Alexandra Butoi and Ryan Cotterell and Anej Sveté

- **Tutorial 5– Watermarking for Large Language Model**

Room: Lotus 5-7 (Level 22)

Xuandong Zhao, and Yu-Xiang Wang, and Lei Li

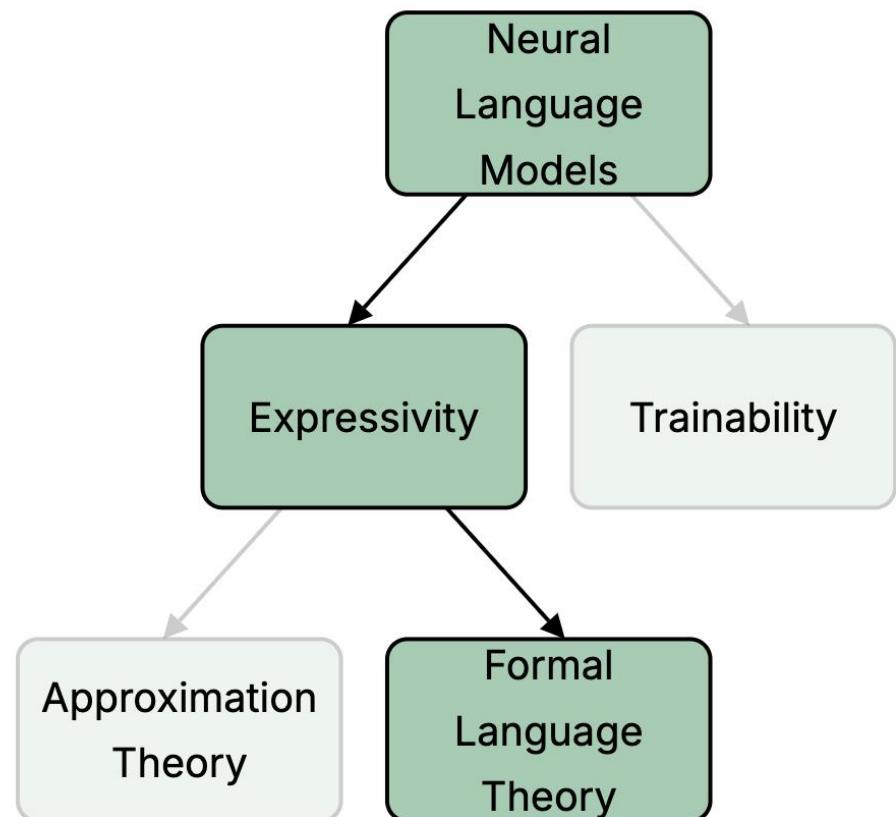
- **Tutorial 6; Presentation Matters: How to Communicate Science in the NLP Venues and in the Wild?**

Room: World Ballroom B (Level 23)

Sarvnaz Karimi, and Cecile Paris, and Gholamreza Haffari

T4: Representational Capacity of Neural Language Models

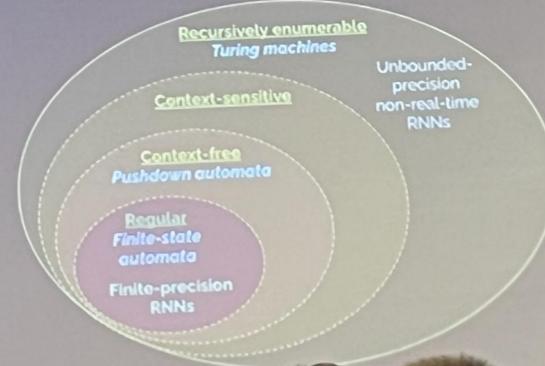
- A group of ETH Zürich led by Ryan Cotterell
- formal language theory to understand the representational capacity of LLMs
- RNN vs. Finite-State Automata
- LLMs vs. finite-state automata and Turing machines.



Our Second Theorem of the Day

- The intuitive connection between RNNs and finite-state automata can be very concretely formalized

Theorem: Finite-precision real-time RNNs are equivalent to FSAs.



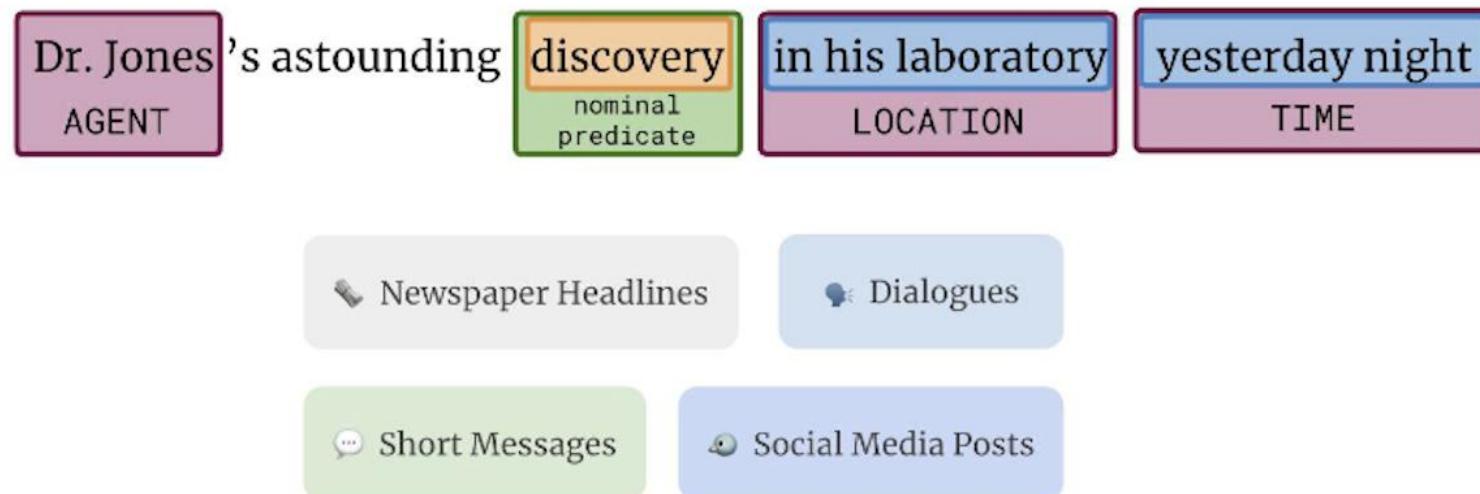
Semantics主题 (oral)

- - NounAtlas: Filling the Gap in Nominal Semantic Role Labeling
Roberto Navigli, Marco Lo Pinto, Pasquale Silvestri, Dennis Rotondi, Simone Ciciliano, Alessandro Scirè
- - UG-schematic Annotation for Event Nominals: A Case Study in Mandarin Chinese
Wenxi Li | Yutong Zhang | Guy Emerson | Weiwei Sun
Computational Linguistics, Volume 50, Issue 2 - June 2023
- Distributional Inclusion Hypothesis and Quantifications:
Probing for Hypernymy in Functional Distributional Semantics
Chun Hei Lo, Wai Lam, Hong Cheng, Guy Emerson

NounAtlas: Filling the Gap in Nominal Semantic Role Labeling

Nominal SRL is under studied

Existing SRL has been primarily focused on verbal predicates.
But nominal predicates are frequent in real-world settings!



Inventory

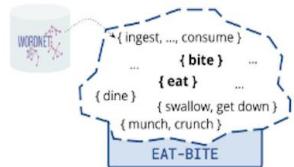
NounAtlas

Filling the gap in nominal SRL with NounAtlas!

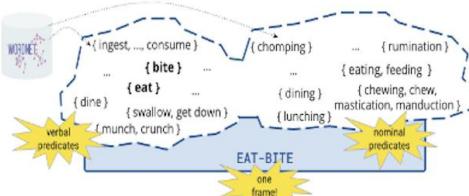
Creating a huge inventory for Nominal SRL

Expanding VerbAtlas frames

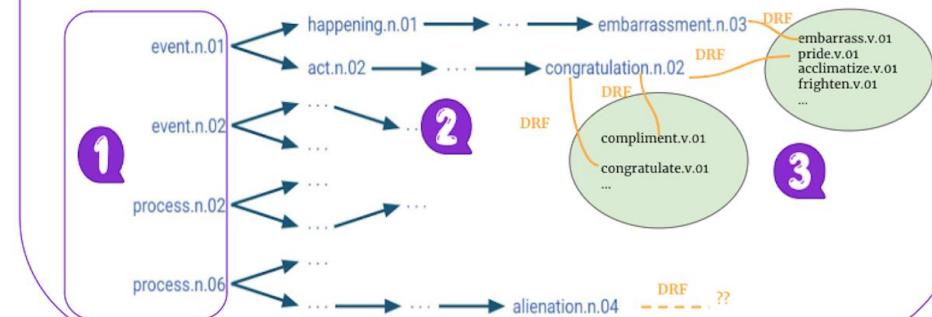
We start from the existing **VerbAtlas frames**, which groups together semantically-related verbal predicates.



We expand the VerbAtlas frames by adding **WordNet nominal synsets** that convey similar semantics.

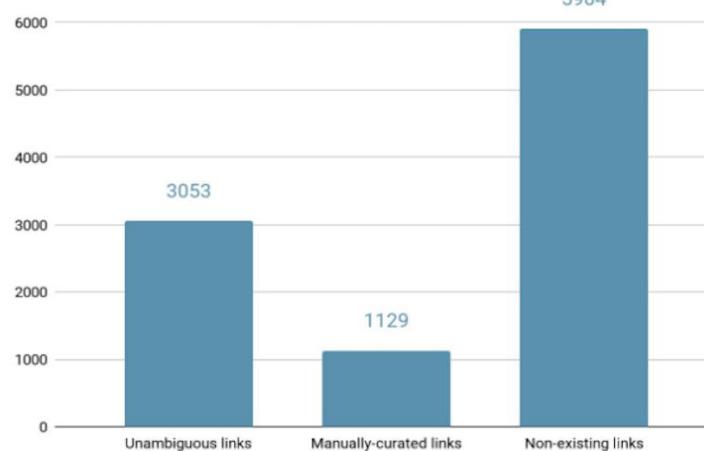


- 1 We manually selected the most prominent **nominal synsets** from WordNet
- 2 We collect all their **nominal descendants** (i.e. via direct/indirect WordNet hyponym relations)
- 3 We link nominal synsets to their verbal counterparts through WordNet's **Derivationally Related Forms (DRF)** edges



Evaluation

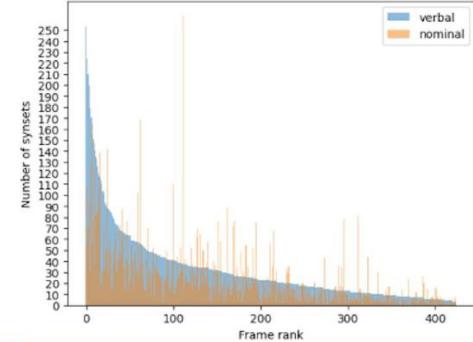
1. Unambiguous links, evaluated on 100 items: 82% in agreement with manual annotations, 18% made up of equally-valid selections
2. Ambiguous links, evaluated on 100 items: tie broken through manual curation, Cohen's Kappa = 0.57 (moderate)



3. Missing links: We use a Cross-Encoder trained on the other types of link to score predicates and, aggregated on ≤ 10 predicates, frames. We consider the top-5 frames and manually link.



- As a result of manual annotation, 99.6% of Unlinked synsets were added to a frame
- 88.2% of correct frames contained in our model's top-5 ranking, 49.3% in the top-1



Corpus + model

Creating a nominal SRL corpus

Predicate nominalization: Generation

Idea → start from SemCor (Miller et al. 1993):

Verbal predicates in SemCor annotated with WordNet synsets are directly linked to **VerbAtlas frames!**

- We retain all SemCor sentences featuring **verbal predicates**
- We generate the corresponding nominalized version of each sentence using **Gemini-Pro**

Change the sentence by nominalizing the verb "married" indicated by
* Use exactly one of these deverbal nouns: "marriage ceremony",
"marriage", "wedding". Indicate the chosen deverbal noun with **:
"The couple was **married** Aug. 2, 1913.

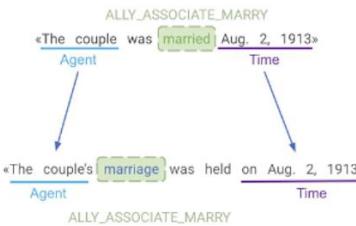
◆ The couple's **marriage** was held on Aug. 2, 1913.
□ ○ ♀

- We validate the generation by keeping the sentence if:

1. The deverbal noun is **identified**
2. The deverbal noun is a **noun**
3. Its **lemma** is among the candidates provided

Verbal-to-nominal role propagation

We annotate the verbal sentences with InVeRo-XL (Conia et al. 2021)



We validate ~500 random sentences to generate a **gold test set** for Nominal SRL:

unchanged sentences: **92.11%**
unchanged frames: **95.56%**
unchanged role spans: **77.11%**
unchanged role labels: **95.27%**



A unified model for SRL

Setup

Backbone model: RoBERTa-based model from Conia and Navigli (2020)

Training data:

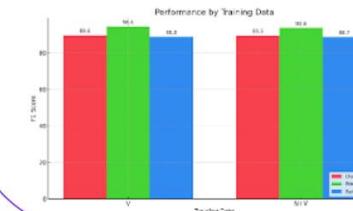
- Nominal SRL: Our dataset
- Verbal SRL: OntoNotes 5.0 (V)
- The combination of the two datasets (N+V)

Test data:

- Our ~500 manually-curated nominal sentences (nouns)
- The OntoNotes test set (verbs) converted to VerbAtlas annotations



OntoNotes verbal test set



Our model, **jointly-trained on nominal and verbal data (N+V)**, achieves competitive performance on nominal SRL!

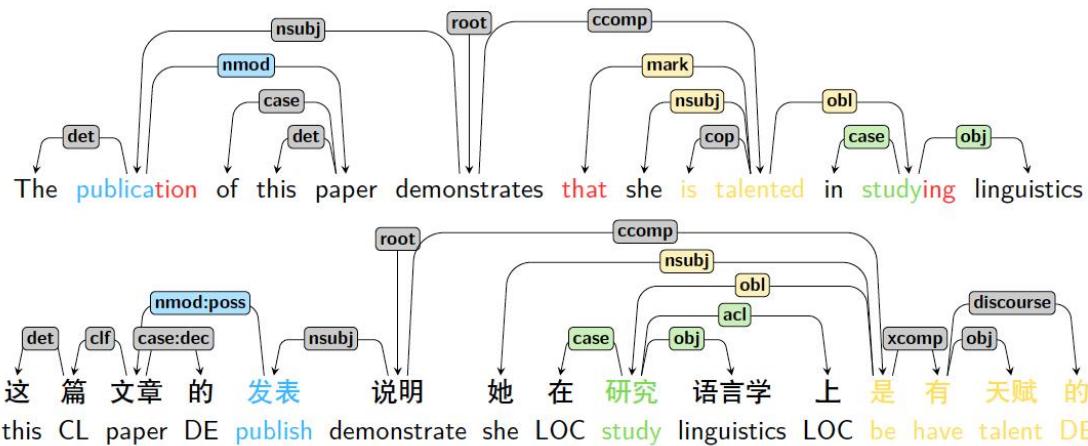


Also robust on the OntoNotes test set (verb frames).

Thanks to NounAtlas, we devised a unified approach for nominal and verbal SRL!

Background

Multilingual Heterogeneity of Event Nominals



- **Identification:** Languages differ in their use of functional morphemes for nominalizations and the types of morphemes they use.
- **Nominal SRL:** Languages differ in how they realize participants of event nominals through syntactic constructions

Research Question

Multilingual Heterogeneity of Event Nominals

	Mandarin	English	German
Subject with nominative case	-	?	+
Accusative case on object	-	?	+
Projection of outer Aspect	-	+	+
Modal or auxiliary verb	+	+	+
Complementizer	-	+	+
Verb suffix	-	+	+
Genitive/PP-subject	-	+	+
Genitive/PP-object	-	+	+
Gender	-	?	+
Quantity	-	+	+
Determiner	-	+	+
Noun suffix	-	+	+

Table: Verb- and noun-related features of event nominals

Research Question

How can we accommodate multilingually heterogeneous phenomena in a unified way, or more theoretically, uncover universals of the world's languages beneath their surface-level variations?

Distributional Inclusion Hypothesis and Quantifications: Probing for Hypernymy in Functional Distributional Semantics

DIH and Quantifications

DIH. r_2 is a hypernym of r_1 iff r_1 's characteristic contexts $\subseteq r_2$'s.

Quantifications. A corpus with only universally quantified statements results in the reverse of DIH (rDIH).

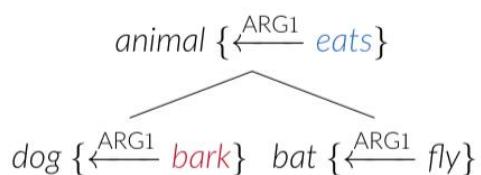


Figure 1. A taxonomic hierarchy of nouns. Next to each noun are the contexts applicable to it and its descendants.

Corpus 1 (DIH)	Corpus 2 (rDIH)
some dog <i>barks</i>	every dog <i>barks</i>
some animal <i>barks</i>	every dog <i>eats</i>
some bat <i>flies</i>	every bat <i>flies</i>
some animal <i>flies</i>	every bat <i>eats</i>
some animal <i>eats</i>	every animal <i>eats</i>

Table 1. Corpora generated from the hierarchy in Fig. 1.

FDS

Entity Vectors. $z \in \mathbb{R}^d$

Truth-Conditional Semantic Functions.

$$\begin{aligned} t^{(dog)}(z) &= P(\text{dog}(z) = \top | z) \\ &= \text{sigmoid}\left(v^{(dog)\top} z + b^{(dog)}\right) \end{aligned}$$

Representing Hypernymy.

$$\forall z \text{ s.t. } \|z\|_2 \leq 1: t^{(dog)}(z) < t^{(animal)}(z)$$

which is true iff $s(\text{dog}, \text{animal}) > 0$, where

$$s(r_h, r_H) = b^{(r_H)} - b^{(r_h)} - \|v^{(r_H)} - v^{(r_h)}\|_2$$

Model Training. By Lo et al. (2023), given a DMRS graph:

$$\text{some} \xrightarrow{\text{RSTR}} \text{dog} \xleftarrow{\text{ARG1}} \text{bark}$$

Variational Inference: $q_\phi(z | \xleftarrow{\text{ARG1}} \text{bark})$

Reconstruction: $\max \ln \mathbb{E}_{q_\phi(z | \xleftarrow{\text{ARG1}} \text{bark})} [t^{(dog)}(z)] + \dots$

New objective (FDS_∀). Optimizes over regions of the entity space for handling universal quantifications

Posters



What does Kiki look like?

Cross-modal associations between speech sounds and visual shapes in vision-and-language models

Tessa Verhoef, Kiana Shahrasbi and Tom Kouwenhoven

Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands
Correspondence: t.verhoef@liacs.leidenuniv.nl, t.kouwenhoven@liacs.leidenuniv.nl

INTRODUCTION

Humans have clear cross-modal preferences when matching novel words to visual shapes.

The development of multimodal models, including VLMs, can potentially revolutionise how machines understand and interact with humans.

Do VLMs associate non-words and visual stimuli in a human-like way?

Bouba or Kiki? From [5,9]

BACKGROUND

Non-arbitrariness as a **general property of language** [8]. Affects language **learning** and shapes language **emergence** [4,9,10].

Possible sources:

- Orthography [2]
- Acoustics and articulation [6,9,11]
- Affective–semantic properties of vocal communication [7]
- Physical properties relating to audiovisual regularities in the environment [3]

Alper and Averbuch-Elor[1] reported strong evidence for a bouba-kiki effect in CLIP and Stable Diffusion ➤ surprising given model training and the absence of relevant data sources such as auditory information and experience with physical object properties.

METHODS

Models	Architecture	Attention	#Params	#img, caps (M)
CLIP	Diffusion Model	None	5.1B	400, 4102
GPT-4o	Specified	None	~	~
VLT	Single stream	Merged	97.4M	4, 10, 9, 83
BLIP-2	Double stream	Q-Holes	3.8B	129, 258
BLIP-4o	Unknown	Unknown	Unknown	Unknown

Experiments: Probed VLMs using a cognitive science paradigm to test for the bouba-kiki effect

Images:

Originals from [5,9] as well as original image pairs from prior human experiments and entirely novel generated images, following [7].

Examples ↪

Pseudowords:

from [7]

Sonorant consonants:	m, n, l
Rounded vowels:	oo, oh, ah
Non-rounded vowels:	ee, uh, ay
Examples:	monoooh, loohm, marmoo
Examples:	teekay, kuhpeh, keepay

4 separate tests: Pseudoword generation (PWG) and probability score comparison (PSC), each with single syllables (1) or two-syllable words (2).

RESULTS

Fig 1: Curved or Jagged images matched to Sonorant consonants (left bars) or Rounded vowel (right bars) in PWG1 test. The expected pattern would show higher bars for the Curved than for the Jagged shapes in both sets. Only true for CLIP and GPT-4o.

Fig 2: Probability scores in PWG1 for two pairs of original pseudowords (Bouba & Kiki), taken & measured with 4 generated syllable types: Sonorant-Rounded (S-R, expected to be most Curved), Sonorant-Sonorant (S-S, Pseudo-Rounded (P-R) and Pseudo-Sonorant (P-S, expected to be most Jagged)), tested with Curved or Jagged shapes. No effect found.

DISCUSSION

It is too early to conclude that VLMs understand sound-symbolism or map visuo-linguistic representations in a human-like way. Results depend heavily on which specific model is tested and how the task is formulated. However, results tentatively suggest that cross-modal preferences can, to some extent, be learned from statistical regularities in data.

Model features such as architecture/design, training objective, number of parameters, and input data seem to affect the results.

These findings inform discussions on the origins of the bouba-kiki effect in human cognition and future developments of VLMs that align well with human cross-modal associations.

REFERENCES

1. Alper, M., and Averbuch-Elor, D. 2023. Evidence for cross-modal associations in human-like ways. In Proc. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 1–10. CVPR ’23, New Orleans, LA, USA, June 2023. The work presented in this paper has not been peer-reviewed and is posted in arXiv as a preprint. It may change before its presentation at the conference.
2. Baudot, J.-P., and J. C. Chastang. 1996. The Bouba-Kiki effect: a case of semantic priming. Psychol. Res. 59, 1: 1–10.
3. Baudot, J.-P., and J. C. Chastang. 1998. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 62, 1: 1–10.
4. Chastang, J. C., and J.-P. Baudot. 1996. The Bouba-Kiki effect: a case of semantic priming. Psychol. Res. 59, 1: 1–10.
5. Chastang, J. C., and J.-P. Baudot. 1998. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 62, 1: 1–10.
6. Chastang, J. C., and J.-P. Baudot. 2000. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 64, 1: 1–10.
7. Chastang, J. C., and J.-P. Baudot. 2002. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 66, 1: 1–10.
8. Chastang, J. C., and J.-P. Baudot. 2004. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 68, 1: 1–10.
9. Chastang, J. C., and J.-P. Baudot. 2006. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 70, 1: 1–10.
10. Chastang, J. C., and J.-P. Baudot. 2008. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 72, 1: 1–10.
11. Chastang, J. C., and J.-P. Baudot. 2010. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 74, 1: 1–10.
12. Chastang, J. C., and J.-P. Baudot. 2012. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 76, 1: 1–10.
13. Chastang, J. C., and J.-P. Baudot. 2014. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 78, 1: 1–10.
14. Chastang, J. C., and J.-P. Baudot. 2016. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 80, 1: 1–10.
15. Chastang, J. C., and J.-P. Baudot. 2018. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 82, 1: 1–10.
16. Chastang, J. C., and J.-P. Baudot. 2020. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 84, 1: 1–10.
17. Chastang, J. C., and J.-P. Baudot. 2022. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 86, 1: 1–10.
18. Chastang, J. C., and J.-P. Baudot. 2024. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 90, 1: 1–10.
19. Chastang, J. C., and J.-P. Baudot. 2026. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 94, 1: 1–10.
20. Chastang, J. C., and J.-P. Baudot. 2028. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 98, 1: 1–10.
21. Chastang, J. C., and J.-P. Baudot. 2030. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 102, 1: 1–10.
22. Chastang, J. C., and J.-P. Baudot. 2032. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 104, 1: 1–10.
23. Chastang, J. C., and J.-P. Baudot. 2034. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 106, 1: 1–10.
24. Chastang, J. C., and J.-P. Baudot. 2036. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 108, 1: 1–10.
25. Chastang, J. C., and J.-P. Baudot. 2038. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 110, 1: 1–10.
26. Chastang, J. C., and J.-P. Baudot. 2040. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 112, 1: 1–10.
27. Chastang, J. C., and J.-P. Baudot. 2042. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 114, 1: 1–10.
28. Chastang, J. C., and J.-P. Baudot. 2044. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 116, 1: 1–10.
29. Chastang, J. C., and J.-P. Baudot. 2046. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 118, 1: 1–10.
30. Chastang, J. C., and J.-P. Baudot. 2048. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 120, 1: 1–10.
31. Chastang, J. C., and J.-P. Baudot. 2050. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 122, 1: 1–10.
32. Chastang, J. C., and J.-P. Baudot. 2052. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 124, 1: 1–10.
33. Chastang, J. C., and J.-P. Baudot. 2054. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 126, 1: 1–10.
34. Chastang, J. C., and J.-P. Baudot. 2056. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 128, 1: 1–10.
35. Chastang, J. C., and J.-P. Baudot. 2058. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 130, 1: 1–10.
36. Chastang, J. C., and J.-P. Baudot. 2060. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 132, 1: 1–10.
37. Chastang, J. C., and J.-P. Baudot. 2062. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 134, 1: 1–10.
38. Chastang, J. C., and J.-P. Baudot. 2064. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 136, 1: 1–10.
39. Chastang, J. C., and J.-P. Baudot. 2066. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 138, 1: 1–10.
40. Chastang, J. C., and J.-P. Baudot. 2068. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 140, 1: 1–10.
41. Chastang, J. C., and J.-P. Baudot. 2070. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 142, 1: 1–10.
42. Chastang, J. C., and J.-P. Baudot. 2072. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 144, 1: 1–10.
43. Chastang, J. C., and J.-P. Baudot. 2074. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 146, 1: 1–10.
44. Chastang, J. C., and J.-P. Baudot. 2076. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 148, 1: 1–10.
45. Chastang, J. C., and J.-P. Baudot. 2078. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 150, 1: 1–10.
46. Chastang, J. C., and J.-P. Baudot. 2080. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 152, 1: 1–10.
47. Chastang, J. C., and J.-P. Baudot. 2082. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 154, 1: 1–10.
48. Chastang, J. C., and J.-P. Baudot. 2084. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 156, 1: 1–10.
49. Chastang, J. C., and J.-P. Baudot. 2086. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 158, 1: 1–10.
50. Chastang, J. C., and J.-P. Baudot. 2088. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 160, 1: 1–10.
51. Chastang, J. C., and J.-P. Baudot. 2090. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 162, 1: 1–10.
52. Chastang, J. C., and J.-P. Baudot. 2092. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 164, 1: 1–10.
53. Chastang, J. C., and J.-P. Baudot. 2094. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 166, 1: 1–10.
54. Chastang, J. C., and J.-P. Baudot. 2096. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 168, 1: 1–10.
55. Chastang, J. C., and J.-P. Baudot. 2098. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 170, 1: 1–10.
56. Chastang, J. C., and J.-P. Baudot. 2100. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 172, 1: 1–10.
57. Chastang, J. C., and J.-P. Baudot. 2102. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 174, 1: 1–10.
58. Chastang, J. C., and J.-P. Baudot. 2104. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 176, 1: 1–10.
59. Chastang, J. C., and J.-P. Baudot. 2106. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 178, 1: 1–10.
60. Chastang, J. C., and J.-P. Baudot. 2108. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 180, 1: 1–10.
61. Chastang, J. C., and J.-P. Baudot. 2110. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 182, 1: 1–10.
62. Chastang, J. C., and J.-P. Baudot. 2112. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 184, 1: 1–10.
63. Chastang, J. C., and J.-P. Baudot. 2114. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 186, 1: 1–10.
64. Chastang, J. C., and J.-P. Baudot. 2116. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 188, 1: 1–10.
65. Chastang, J. C., and J.-P. Baudot. 2118. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 190, 1: 1–10.
66. Chastang, J. C., and J.-P. Baudot. 2120. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 192, 1: 1–10.
67. Chastang, J. C., and J.-P. Baudot. 2122. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 194, 1: 1–10.
68. Chastang, J. C., and J.-P. Baudot. 2124. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 196, 1: 1–10.
69. Chastang, J. C., and J.-P. Baudot. 2126. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 198, 1: 1–10.
70. Chastang, J. C., and J.-P. Baudot. 2128. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 200, 1: 1–10.
71. Chastang, J. C., and J.-P. Baudot. 2130. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 202, 1: 1–10.
72. Chastang, J. C., and J.-P. Baudot. 2132. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 204, 1: 1–10.
73. Chastang, J. C., and J.-P. Baudot. 2134. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 206, 1: 1–10.
74. Chastang, J. C., and J.-P. Baudot. 2136. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 208, 1: 1–10.
75. Chastang, J. C., and J.-P. Baudot. 2138. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 210, 1: 1–10.
76. Chastang, J. C., and J.-P. Baudot. 2140. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 212, 1: 1–10.
77. Chastang, J. C., and J.-P. Baudot. 2142. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 214, 1: 1–10.
78. Chastang, J. C., and J.-P. Baudot. 2144. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 216, 1: 1–10.
79. Chastang, J. C., and J.-P. Baudot. 2146. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 218, 1: 1–10.
80. Chastang, J. C., and J.-P. Baudot. 2148. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 220, 1: 1–10.
81. Chastang, J. C., and J.-P. Baudot. 2150. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 222, 1: 1–10.
82. Chastang, J. C., and J.-P. Baudot. 2152. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 224, 1: 1–10.
83. Chastang, J. C., and J.-P. Baudot. 2154. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 226, 1: 1–10.
84. Chastang, J. C., and J.-P. Baudot. 2156. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 228, 1: 1–10.
85. Chastang, J. C., and J.-P. Baudot. 2158. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 230, 1: 1–10.
86. Chastang, J. C., and J.-P. Baudot. 2160. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 232, 1: 1–10.
87. Chastang, J. C., and J.-P. Baudot. 2162. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 234, 1: 1–10.
88. Chastang, J. C., and J.-P. Baudot. 2164. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 236, 1: 1–10.
89. Chastang, J. C., and J.-P. Baudot. 2166. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 238, 1: 1–10.
90. Chastang, J. C., and J.-P. Baudot. 2168. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 240, 1: 1–10.
91. Chastang, J. C., and J.-P. Baudot. 2170. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 242, 1: 1–10.
92. Chastang, J. C., and J.-P. Baudot. 2172. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 244, 1: 1–10.
93. Chastang, J. C., and J.-P. Baudot. 2174. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 246, 1: 1–10.
94. Chastang, J. C., and J.-P. Baudot. 2176. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 248, 1: 1–10.
95. Chastang, J. C., and J.-P. Baudot. 2178. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 250, 1: 1–10.
96. Chastang, J. C., and J.-P. Baudot. 2180. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 252, 1: 1–10.
97. Chastang, J. C., and J.-P. Baudot. 2182. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 254, 1: 1–10.
98. Chastang, J. C., and J.-P. Baudot. 2184. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 256, 1: 1–10.
99. Chastang, J. C., and J.-P. Baudot. 2186. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 258, 1: 1–10.
100. Chastang, J. C., and J.-P. Baudot. 2188. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 260, 1: 1–10.
101. Chastang, J. C., and J.-P. Baudot. 2190. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 262, 1: 1–10.
102. Chastang, J. C., and J.-P. Baudot. 2192. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 264, 1: 1–10.
103. Chastang, J. C., and J.-P. Baudot. 2194. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 266, 1: 1–10.
104. Chastang, J. C., and J.-P. Baudot. 2196. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 268, 1: 1–10.
105. Chastang, J. C., and J.-P. Baudot. 2198. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 270, 1: 1–10.
106. Chastang, J. C., and J.-P. Baudot. 2200. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 272, 1: 1–10.
107. Chastang, J. C., and J.-P. Baudot. 2202. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 274, 1: 1–10.
108. Chastang, J. C., and J.-P. Baudot. 2204. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 276, 1: 1–10.
109. Chastang, J. C., and J.-P. Baudot. 2206. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 278, 1: 1–10.
110. Chastang, J. C., and J.-P. Baudot. 2208. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 280, 1: 1–10.
111. Chastang, J. C., and J.-P. Baudot. 2210. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 282, 1: 1–10.
112. Chastang, J. C., and J.-P. Baudot. 2212. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 284, 1: 1–10.
113. Chastang, J. C., and J.-P. Baudot. 2214. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 286, 1: 1–10.
114. Chastang, J. C., and J.-P. Baudot. 2216. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 288, 1: 1–10.
115. Chastang, J. C., and J.-P. Baudot. 2218. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 290, 1: 1–10.
116. Chastang, J. C., and J.-P. Baudot. 2220. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 292, 1: 1–10.
117. Chastang, J. C., and J.-P. Baudot. 2222. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 294, 1: 1–10.
118. Chastang, J. C., and J.-P. Baudot. 2224. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 296, 1: 1–10.
119. Chastang, J. C., and J.-P. Baudot. 2226. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 298, 1: 1–10.
120. Chastang, J. C., and J.-P. Baudot. 2228. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 300, 1: 1–10.
121. Chastang, J. C., and J.-P. Baudot. 2230. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 302, 1: 1–10.
122. Chastang, J. C., and J.-P. Baudot. 2232. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 304, 1: 1–10.
123. Chastang, J. C., and J.-P. Baudot. 2234. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 306, 1: 1–10.
124. Chastang, J. C., and J.-P. Baudot. 2236. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 308, 1: 1–10.
125. Chastang, J. C., and J.-P. Baudot. 2238. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 310, 1: 1–10.
126. Chastang, J. C., and J.-P. Baudot. 2240. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 312, 1: 1–10.
127. Chastang, J. C., and J.-P. Baudot. 2242. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 314, 1: 1–10.
128. Chastang, J. C., and J.-P. Baudot. 2244. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 316, 1: 1–10.
129. Chastang, J. C., and J.-P. Baudot. 2246. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 318, 1: 1–10.
130. Chastang, J. C., and J.-P. Baudot. 2248. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 320, 1: 1–10.
131. Chastang, J. C., and J.-P. Baudot. 2250. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 322, 1: 1–10.
132. Chastang, J. C., and J.-P. Baudot. 2252. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 324, 1: 1–10.
133. Chastang, J. C., and J.-P. Baudot. 2254. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 326, 1: 1–10.
134. Chastang, J. C., and J.-P. Baudot. 2256. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 328, 1: 1–10.
135. Chastang, J. C., and J.-P. Baudot. 2258. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 330, 1: 1–10.
136. Chastang, J. C., and J.-P. Baudot. 2260. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 332, 1: 1–10.
137. Chastang, J. C., and J.-P. Baudot. 2262. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 334, 1: 1–10.
138. Chastang, J. C., and J.-P. Baudot. 2264. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 336, 1: 1–10.
139. Chastang, J. C., and J.-P. Baudot. 2266. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 338, 1: 1–10.
140. Chastang, J. C., and J.-P. Baudot. 2268. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 340, 1: 1–10.
141. Chastang, J. C., and J.-P. Baudot. 2270. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 342, 1: 1–10.
142. Chastang, J. C., and J.-P. Baudot. 2272. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 344, 1: 1–10.
143. Chastang, J. C., and J.-P. Baudot. 2274. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 346, 1: 1–10.
144. Chastang, J. C., and J.-P. Baudot. 2276. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 348, 1: 1–10.
145. Chastang, J. C., and J.-P. Baudot. 2278. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 350, 1: 1–10.
146. Chastang, J. C., and J.-P. Baudot. 2280. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 352, 1: 1–10.
147. Chastang, J. C., and J.-P. Baudot. 2282. Semantic priming and the Bouba-Kiki effect. Psychol. Res. 354, 1: 1–10.
148. Chastang, J.

Aligning to Adults Is Easy, Aligning to Children Is Hard: A Study of Linguistic Alignment in Dialogue Systems

Dorothea French, Sidney D'Mello, Katharina von der Wense
University of Colorado Boulder

Introduction

Over the course of a conversation people start to align with each other. They use similar words, syntax, and talk about similar ideas. This phenomenon, called linguistic alignment, helps conversational participants understand each other and reduces the effort needed to do so, amongst other positive effects.

Linguistic alignment is particularly important when one of the participants is a child, or a non-fluent speaker of the language. Levels of parent's alignment have been shown to correspond to a child's vocabulary and language development. Repetition and correction also help language learners improve naturally.

In this era, characterized by widespread use of dialogue system, it is important to ask how well they can communicate with humans. Alignment of dialogue systems can reduce frustration and perceived task load. It is also a critical component of communication. Thus we must ask, how well do our systems do? Do they reach human-like levels, or is there an area of friction?

In this work we ask two questions:

To what degree do two state-of-the-art dialogue systems - ChatGPT and Llama2 - align to users, and how does their alignment compare to that typically seen between humans?

To what degree do ChatGPT and Llama2 align to children (i.e. non-fluent speakers), and how does this level of alignment compare to a parent's?



Approach and Data

We use the Switchboard Dialogue Acts Corpus (SWDA) of a series of phone conversations on a variety of topics to assess model alignment with adults.

SWDA Context		CHILDES Context	
A	Any jury's not going to disregard the evidence you know.	MOT	Hm?
B	Uh, that's true	CHI	Where Mommy go?
B	I, I think our judicial system is attorney welfare myself.	MOT	Mommy went to the university this morning to get some books
A	That may very well be.	CHI	Where's Mommy's books?
Response		Response	
B	I, I hold it at the utmost confidentiality.	MOT	They're in the hallway in a big bag
GPT	It's definitely a possibility that needs to be looked into.	GPT	Mommy will bring the books home this evening
Llama2	Yeah, it's like, you know, they're just trying	Llama2	Mommy left her books in the car

Table 1: Last lines of SWDA transcript plus true and generated responses

We look at three types of alignment: syntactic, lexical, and semantic between the last turn of the context and the response and compare these metrics between the models and gold responses

From the datasets we select transcript excerpts of 36 turns - 35 turns of context, and to following utterance.

Our two models, ChatGPT-3.5 and Llama2, are prompted with respond to the 35 turns of context as a participant or parent respectively

Looking at the SWDA responses, we see ChatGPT aligns at human-like levels. Whereas Llama2 exhibits elevated levels of alignment. Upon closer inspection, we see that in the responses, ChatGPT's responses are generally better, scoring an average of 4.37 out of 5 for quality, compared to Llama2's 3.5 out of 5. The ChatGPT's responses are more likely to contain novel or information that drives the conversation forward, confirmed by a lower semantic and lexical alignment.

ChatGPT is more repetitive, and it does a very good job mimicking the stylistic elements of the conversation, but often makes less sense.

When responding to children, both models have lower than human levels of alignment. Like with adults, ChatGPT has better responses, and a much lower proportion of a poor (1-2 out of 5) responses. Once again, Llama2 has more repetitive responses, but higher alignment. In both responses sets we see the smaller Llama2 model consistently has higher alignment.

Table 2: Average alignment of responses to SWDA

Table 3: Average alignment of responses to CHILDES

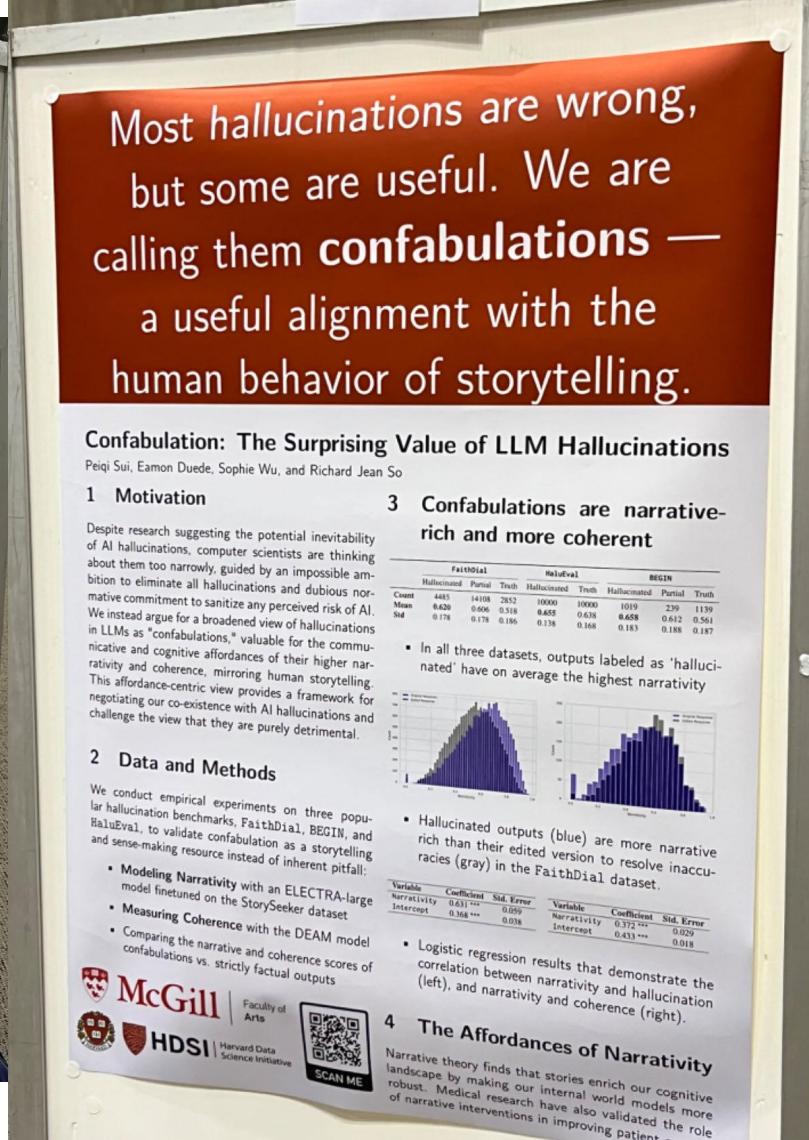
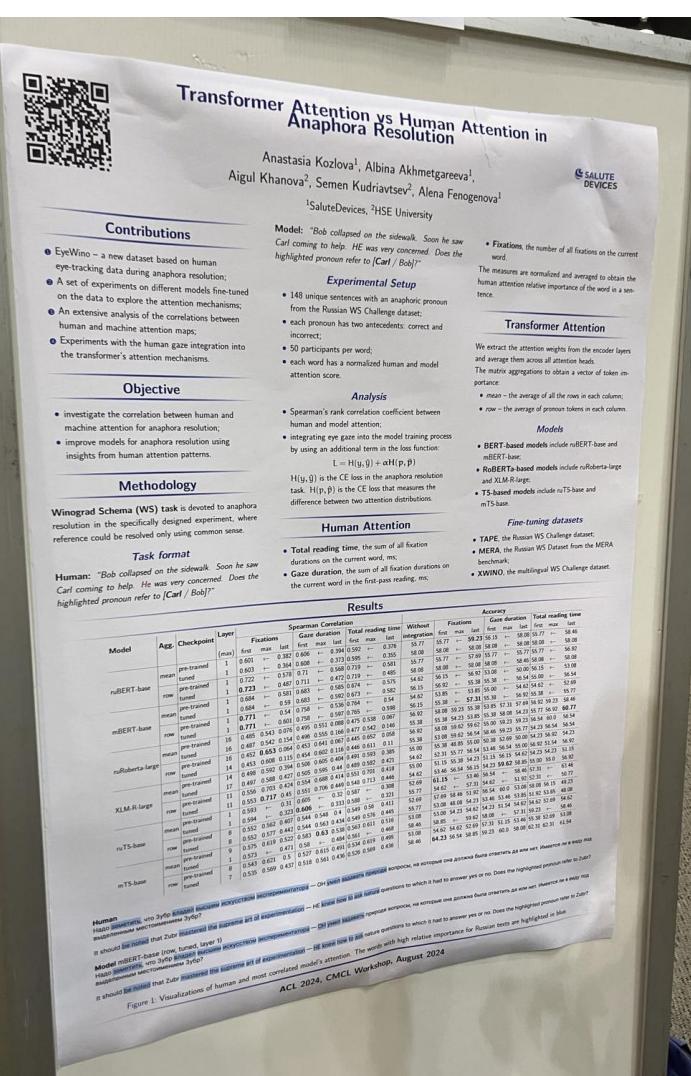
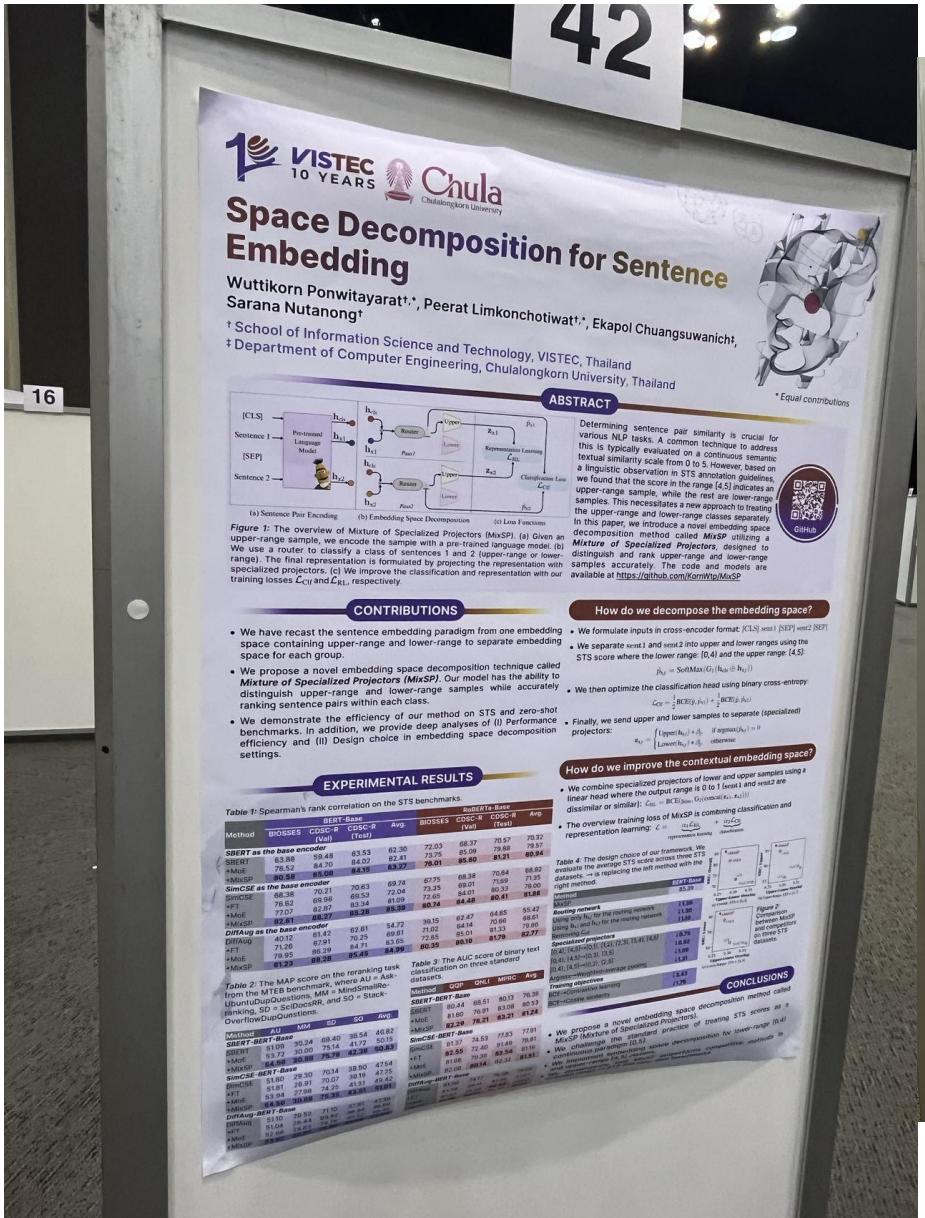
Table 4: Average alignment of responses to CHILDES

Conclusion

Dialogue systems show great potential to assist humans across a variety of tasks. The success of these interactions correlates with linguistic alignment. We find that, when responding to adult speakers, ChatGPT shows approximately human-like alignments and provides constructive responses to Llama2, however, overly mimics the conversation and could be positive when talking with children or language learners as it results in elevated alignment.

We conclude that SOTA dialogue systems can improve in regards to tailoring levels of alignment to match various circumstances, while maintaining quality.

In the future, we plan to investigate alignment to adult learners, non-native speakers, and explore methods to create dialogue systems with a closer alignment to human levels of alignment.



Social Events



颁奖环节

- The 2024 winner of the 1999 Test-of-Time Paper Award is:
- Lillian Lee. 1999. Measures of Distributional Similarity.
- In Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL), pages 25–32, College Park, Maryland, USA.
- The 2024 winner of the 2014 Test-of-Time Paper Award is:
- Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. GloVe: Global Vectors for Word Representation.
- In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1532–1543, Doha, Qatar.

Best Paper Award

- Mission: Impossible Language Models
- Why are Sensitive Functions Hard for Transformers?
- Deciphering Oracle Bone Language with Diffusion Models
- Causal Estimation of Memorisation Profiles
- Aya Model: An Instruction Finetuned Open-Access Multilingual Language Model
- Semisupervised Neural Proto-Language Reconstruction
- Natural Language Satisfiability: Exploring the Problem Distribution and Evaluating Transformer-based Language Models

One more thing... from EMB

ACL is Not an AI Conference

- I am not using “AI” as a synonym for “ML”.
- Machine learning (including deep learning) provides many techniques that are useful for language technology and computational linguistics
 - (Though both of those terms are problematic.)
- The problems of CL/NLP can also be illuminating for questions about ML
- The issues that I am concerned with arise when the focus shifts to “AI”

AI as a research & commercial field

Asks questions like:

- How do we build “thinking machines” that can do “human-like” reasoning?
- How do we build “thinking machines” that can “surpass” humans in cognitive work?
 - (and cure cancer, solve the climate crisis, make end-of-life decisions, etc)
- How do we automate the scientific method?
- How do we automate away such creative work as painting and writing?
 - Or: How do we steal artwork at scale and try to convince people this is “for the common good”?

Compling/NLP asks questions such as

- How are languages similar/different?
- How is information represented in languages?
- How can we build technology that assists with: transcription, translation, summarization, information access ... in different languages?
- How can we evaluate such technology?
- What kinds of intermediate representations are useful for such technology?
- How well do different ML techniques work for different tasks?
- How do language technologies interact with existing systems of power and oppression?

ACL is historically and should remain:

- A venue for people who care about the *language* in language technology
- A community that fosters interdisciplinarity
- A research field that cares about language communities
- ... and, as a result, a space where we can reason about societal impacts of our research and technology

<https://bit.ly/EMB-ACL24>

**Language processing is a
prerequisite for “AI”, but
that doesn’t mean that “AI”
is the only goal of CL/NLP.**

Workshops

- Rep4NLP (Probing -> causal inference)
 - - Latent Space Exploration for Safe and Trustworthy AI by Hassan Sajjad
 - The gap between what language models say and what they know
 - Generalisation in LLMs – and beyond
 - Efficiency as an Inductive Bias: Towards Tokenizer-free and Dynamically Sparse Language Models
- Heng Ji: AI Plays Medicinal Chemist
- Anna Rogers: Ai for reasearch workflow



清华大学
Tsinghua University

Q & A

THANK YOU

Note