ISDS

TU Graz

SCIENCE
PASSION
TECHNOLOGY

Computational Framing Analysis for Polarized Topics Online

PhD Defense: **Markus Reiter-Haas**

2024-06-27

Supervisor: Assoc.Prof. Dipl.-Ing. Dr.techn. Elisabeth Lex
Co-Supervisor: Univ.-Prof. Mag. Dr.rer.soc.oec. Markus Hadler
External Examiner: Prof. Dr. Martin Potthast
Chair: Univ.-Prof. Dipl.-Ing. Dr.techn. Friedrich Fraundorfer

https://iseratho.github.io/thesis

ISDS
INSTITUTE OF INTERACTIVE SYSTEMS AND DATA SCIENCE

Context of the PhD Project

Interdisciplinary research project: Polarization in public opinion

COVID-19 pandemic

Climate change

COVID-19 Vaccine Image by Tim Reckmann from Hamm, Deutschland - SARS-COV-2 Impfstoff, CC BY 2.0, https://commons.wikimedia.org/w/index.php?curid=104894977
Ice Bear Image by Andreas Weith - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=52745369

Talk Outline

1. Motivation of Framing
   - Research Questions
   - Overview of the Background
2. Publications
   - Opinion Polarization
   - Framing Labels
   - Framing Structure
   - Framing Behavior
3. Conclusion
   - Future Research
   - PhD Summary

Framing in Polarized Topics

Framing in **Communication:**
Interpretation due to **Salience**

They prevent the spread

We fight for liberty

COVID-19 pandemic

~~What?~~ → How?

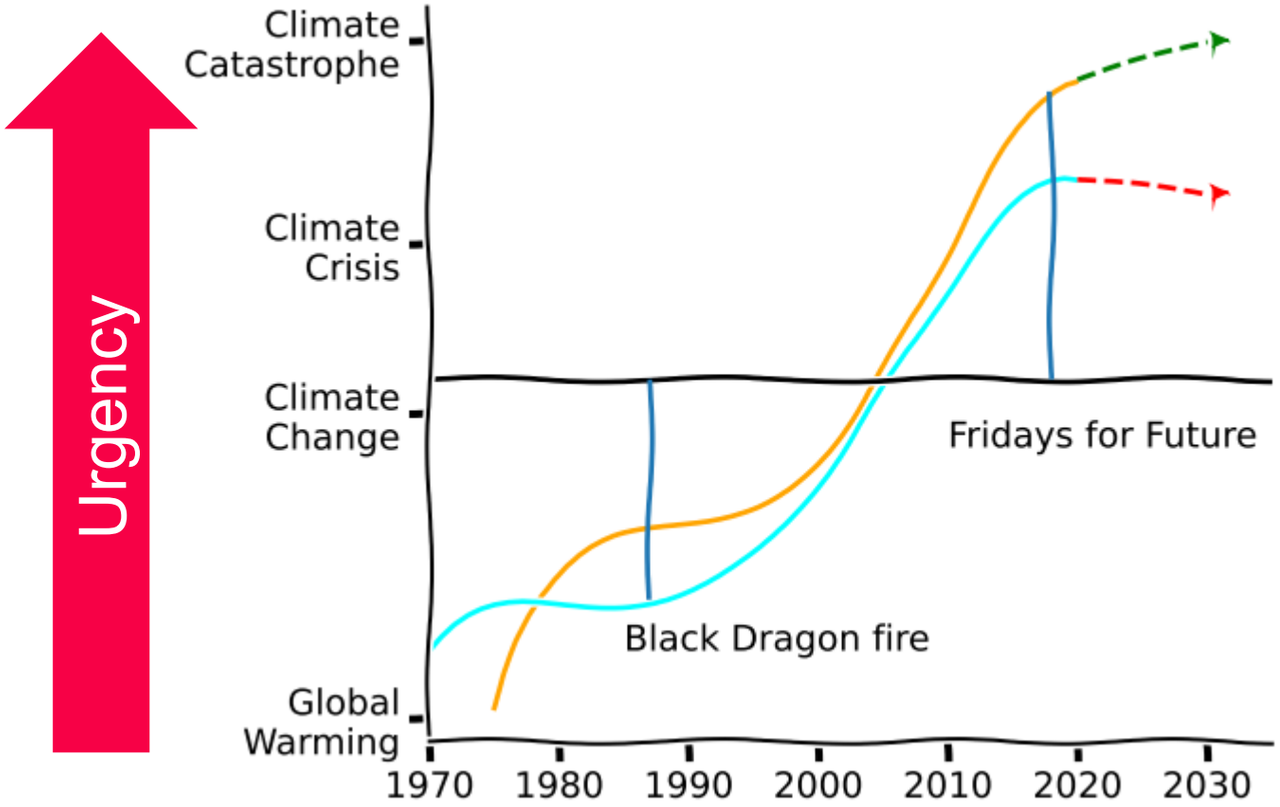
Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. Journal of communication, 43(4), 51-58.

COVID-19 Vaccine Image by Tim Reckmann from Hamm, Deutschland - SARS-COV-2 Impfstoff, CC BY 2.0, https://commons.wikimedia.org/w/index.php?curid=104894977

Complexity of Framing

Framing affects Behavior

Evolution of Climate Change Framing

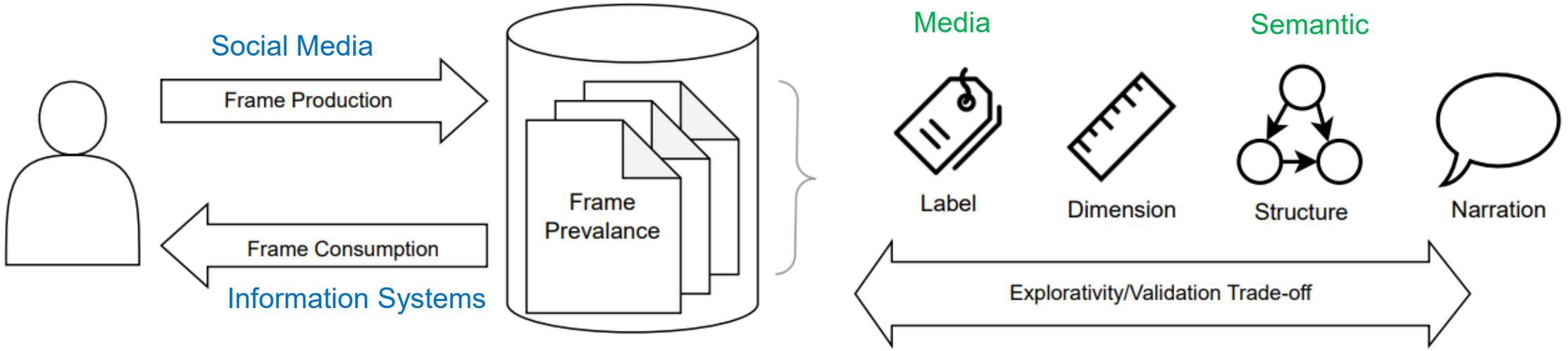


Detection and **Understanding**?

[C8] Reiter-Haas, M., Klösch, B., Hadler, M., & Lex, E. (2024). Computational Narrative Framing: Towards Identifying Frames through Contrasting the Evolution of Narration. In Proceedings of the Text2Story'24 Workshop, Glasgow (Scotland), 24-March-2024.

Ice Bear Image by Andreas Weith - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=52745369

Computational Framing Research



Research Gap

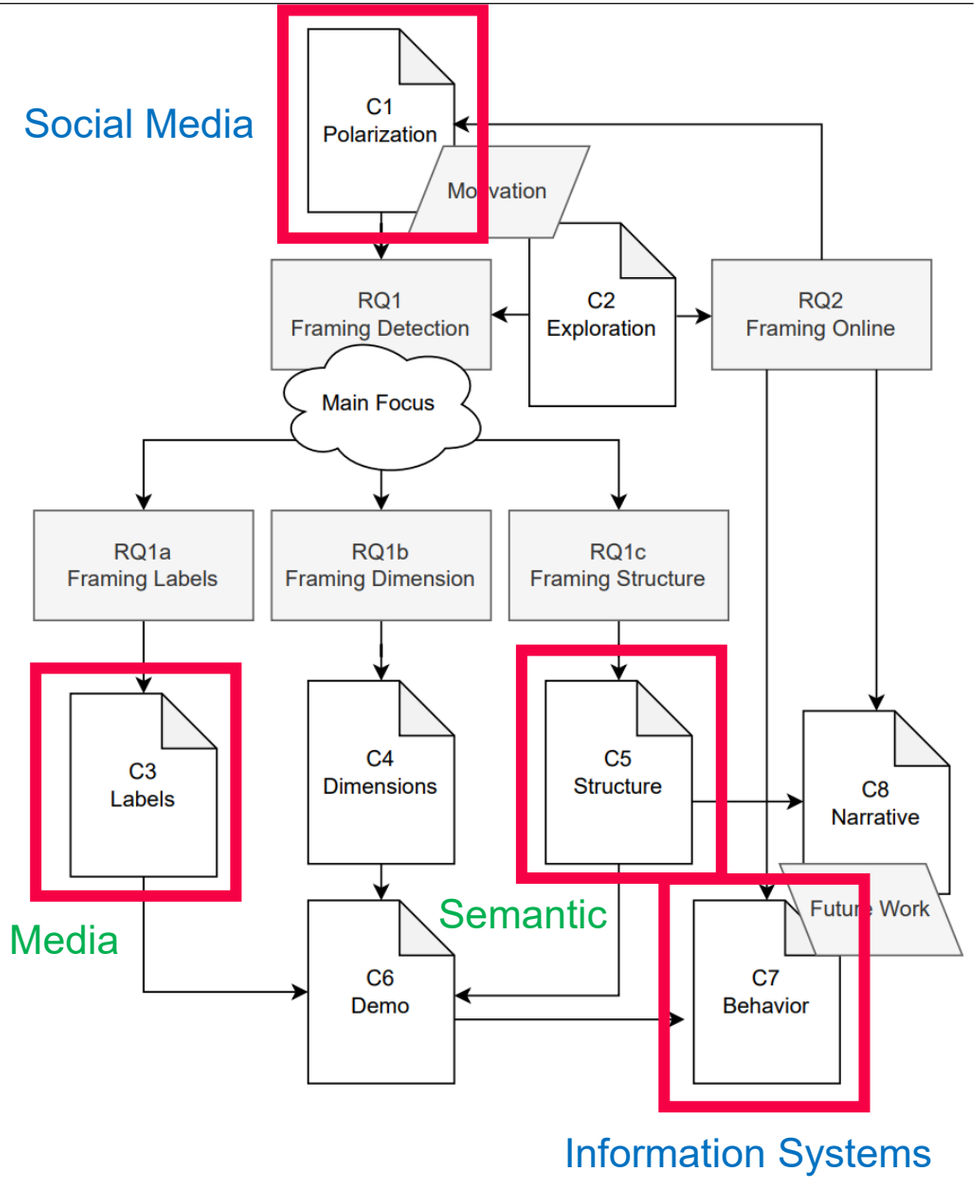
Research Gap

Research Questions

RQ1: How to **detect** differences in the **framing** of online content at various exploratory levels?

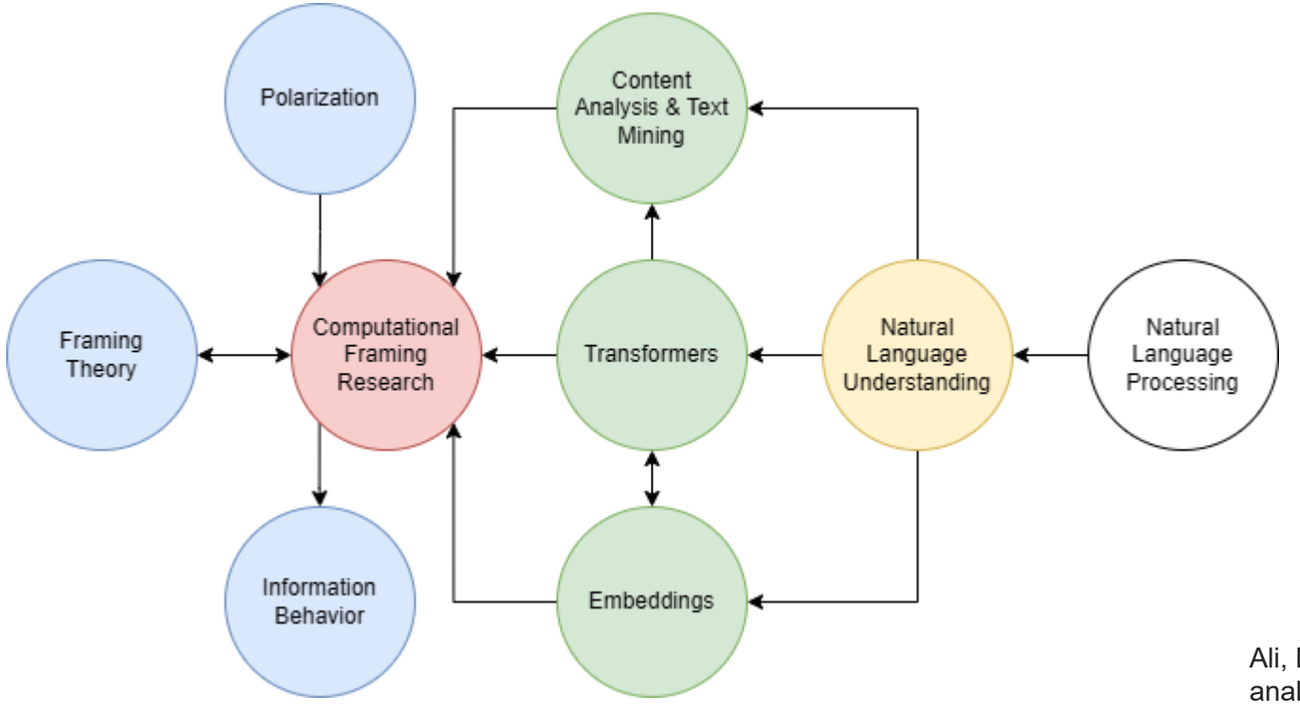
- RQ1a: How to extract **framing labels** with limited annotated data?
- RQ1b: How to extract **framing dimensions** in an unsupervised manner?
- RQ1c: How to extract **framing structure** without prior conceptualization?

RQ2: How does **framing** influence online information **behavior**?



Overview: Interdisciplinary Background

Social Sciences Computer Science



Main Theories

Field of Study

Main Methods

Related Work

Ali, M., & Hassan, N. (2022, December). A survey of computational framing analysis approaches. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing* (pp. 9335-9348).

Talk Outline

1. Motivation of Framing
   - Research Questions
   - Overview of the Background
2. Publications
   - Opinion Polarization
   - Framing Labels
   - Framing Structure
   - Framing Behavior
3. Conclusion
   - Future Research
   - PhD Summary

Polarization: Offline (Survey) vs. Online (Social Media)

Representative quota sample:
Beginning of August 2020 (~1 week)

Rehydrated tweets in German:
January until August 2020

Agreement
1 to 5

Sentiment
-1 to 1

Annotated tweets of users who provided consent and access

Chen, E., Lerman, K., & Ferrara, E. (2020). Tracking social media discourse about the covid-19 pandemic: Development of a public coronavirus twitter data set. JMIR Public Health and Surveillance, 6(2), e19273.

DACH Image by Immanuel Giel, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=15521791
Twitter Logo by Twitter - GitHub, Apache License 2.0,
https://commons.wikimedia.org/w/index.php?curid=112280645

Polarization: Data and Statistics

2560 respondents
(1721 D, 565 A, 274 CH)

25,769
60,218
4,819

79 consented,
20 access → 221 tweets

Guided by bimodality coefficient:

$$\beta = \frac{\gamma^2 + 1}{\kappa + 3\frac{(n-1)^2}{(n-2)(n-3)}}$$

using Skewness γ and Kurtosis κ

DACH Image by Immanuel Giel, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=15521791
Twitter Logo by Twitter - GitHub, Apache License 2.0,
https://commons.wikimedia.org/w/index.php?curid=112280645

Polarization: Congruent Opinions

β		>		&	
	0.67		0.65		0.59
	0.49		0.44		0.44

only 1 discrepancy,
inter-annotator agreement: α = 0.7

Similarity: macro and micro level

Sentiment as proxy ✓

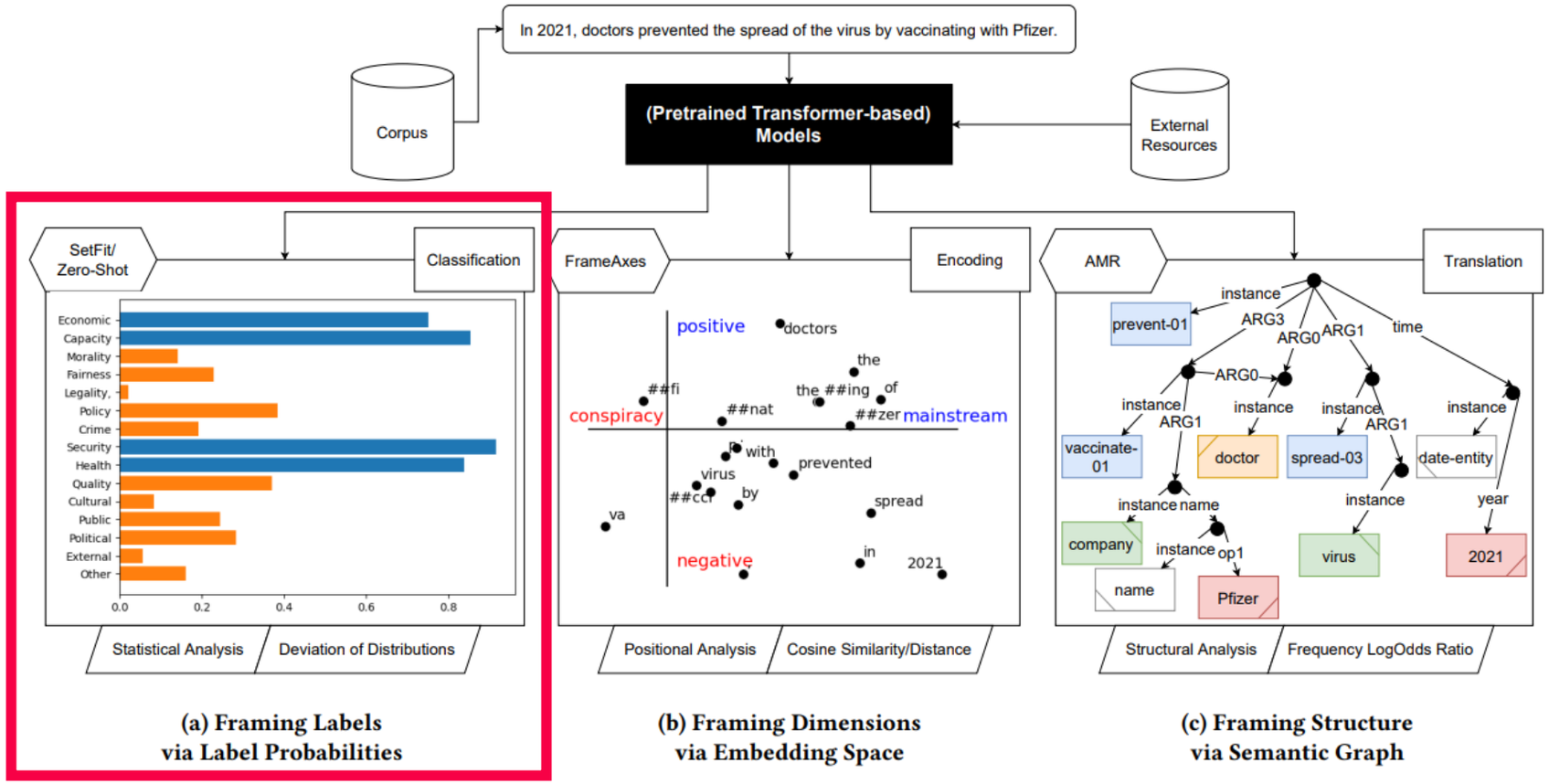
→ Does not capture nuances

[C1] Reiter-Haas, M.*, Klösch, B.*, Hadler, M., & Lex, E. (2023). Polarization of Opinions on COVID-19 Measures: Integrating Twitter and Survey Data. Social Science Computer Review 41 (5), 1811-1835.

DACH Image by Immanuel Giel, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=15521791
Twitter Logo by Twitter - GitHub, Apache License 2.0,
https://commons.wikimedia.org/w/index.php?curid=112280645

RQ1: Framing Detection

RQ1a

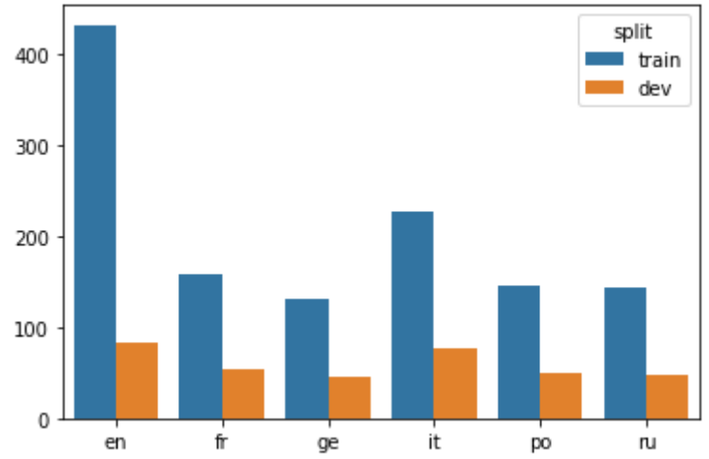


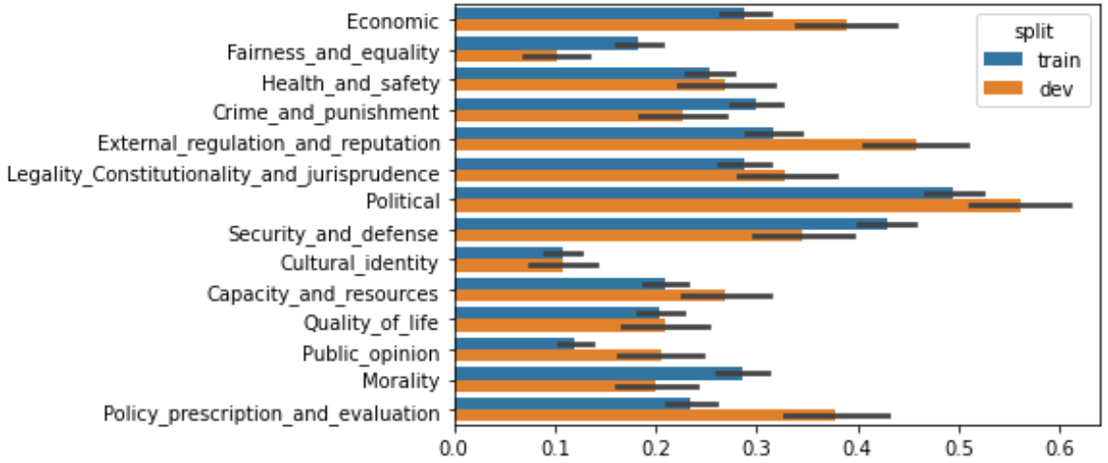
[C2] Reiter-Haas, M. (2023). Exploration of Framing Biases in Polarized Online Content Consumption. In Companion Proceedings of the ACM Web Conference 2023, 560–564. Presented at the Austin, TX, USA.

RQ1a: SemEval 2023 Task 3 Subtask 2

Data sparsity and imbalance

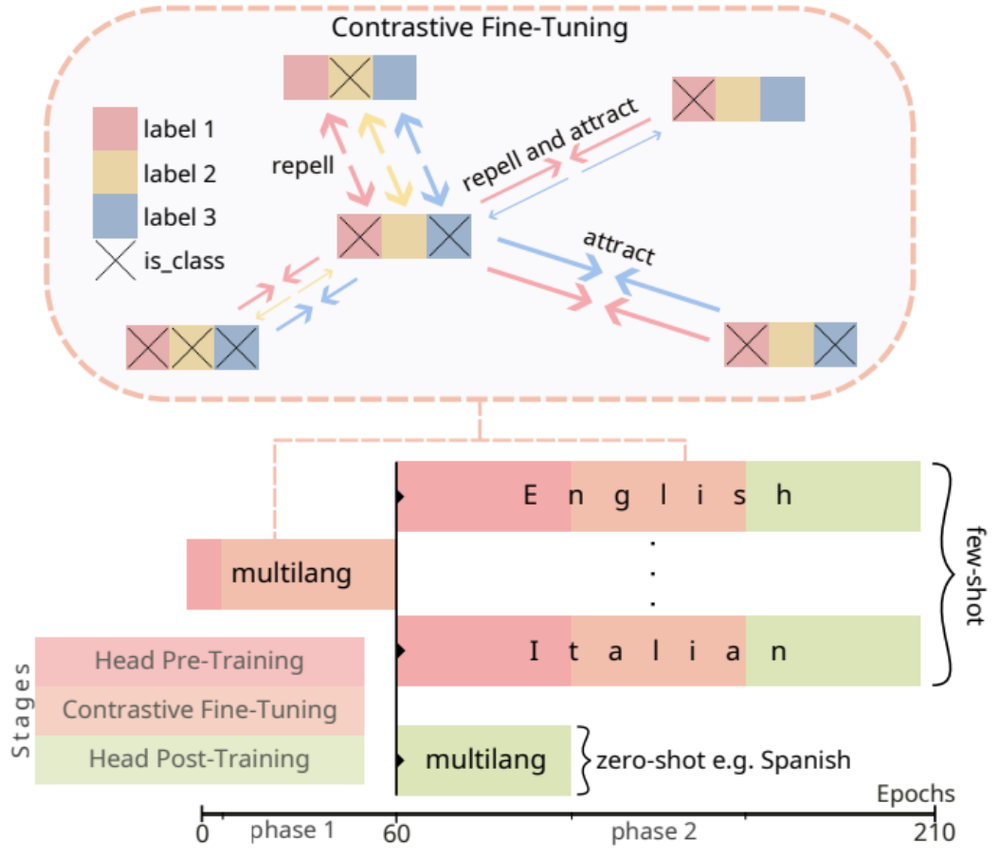
- Multi-class multi-label: 14 Media Frames
- Various polarized topics
  (COVID-19, climate change, abortion, migration, war, …)
- 2049 samples
  (train, dev, test)
- 9 languages
  (6 few-, 3 zero-shot)





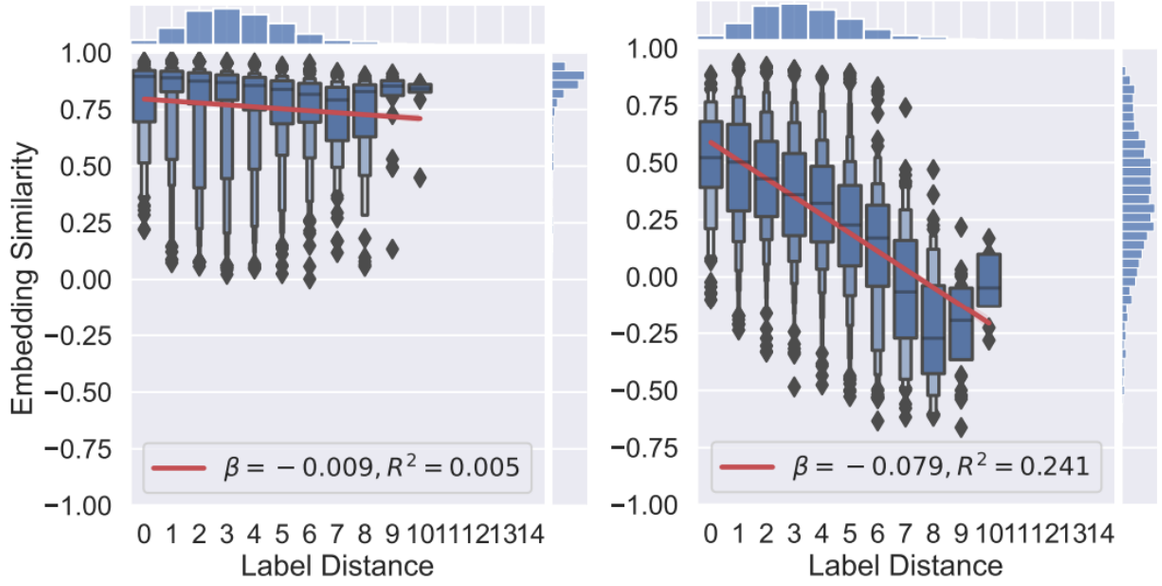
Piskorski, J., Stefanovitch, N., Da San Martino, G., & Nakov, P. (2023, July). Semeval-2023 task 3: Detecting the category, the framing, and the persuasion techniques in online news in a multi-lingual setup. In Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023) (pp. 2343-2361).

RQ1a: Contrastive Pretraining



Our system: **mCPT**

Effects of contrastive pretraining



(b) Before contrastive training

(c) After contrastive training

Zheng, L., Xiong, J., Zhu, Y., & He, J. (2022, August). Contrastive learning with complex heterogeneity. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 2594-2604).

RQ1a: Frame Label Prediction

Improvements over:

- TF-IDF Baseline
- SetFit Library

Knowledge transfer:

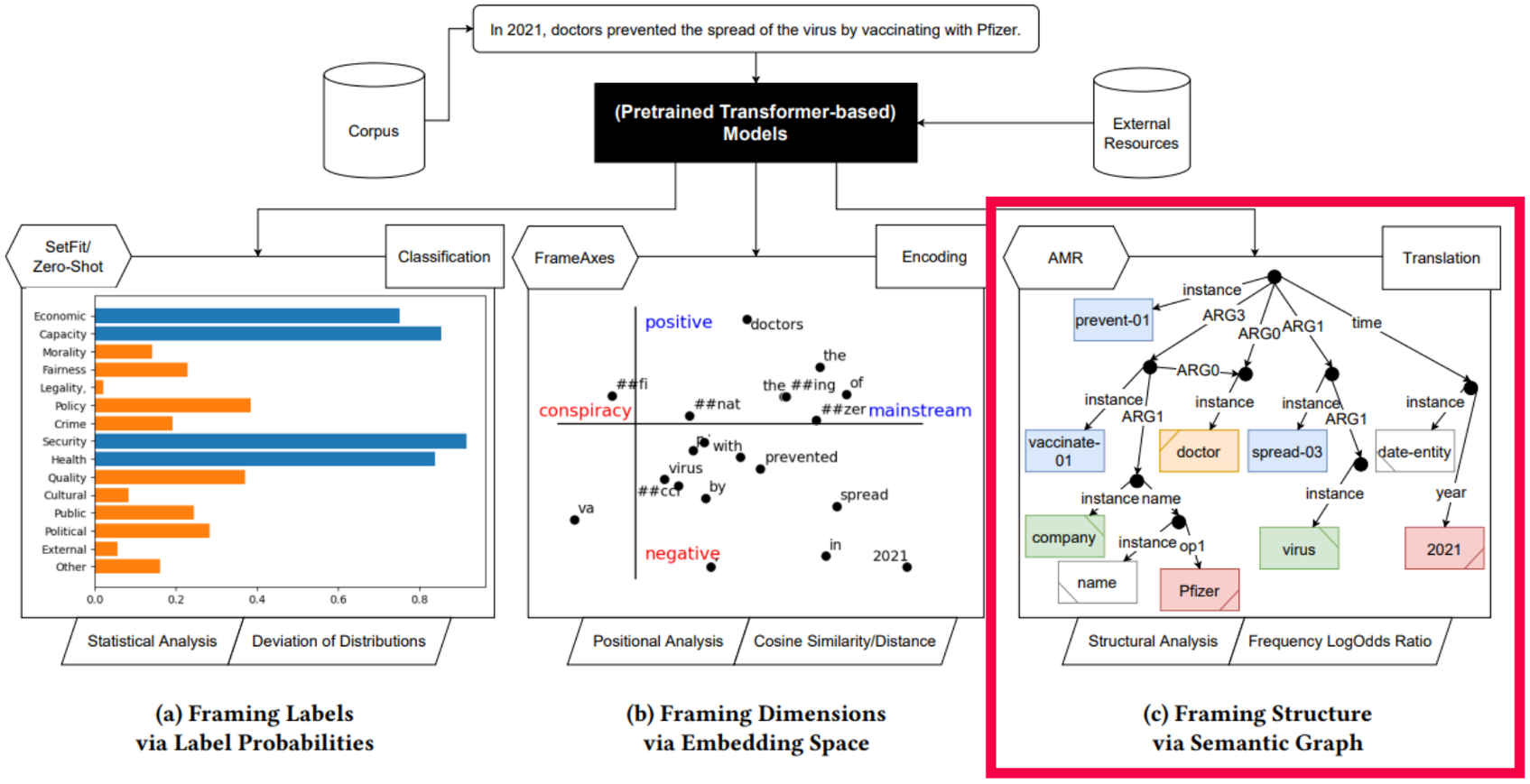
Languages based on Latin script

→ 1st leaderboard position on Spanish

Language	# Samples Train/Dev/Test	Micro-F1 mCPT	Micro-F1 SETFIT	Micro-F1 Base	Macro-F1 mCPT	Macro-F1 SETFIT	Macro-F1 Base	Position #	Position Teams
German ($\mathcal{G}$, L)	132 / 45 / 50	**.622***	.549	.487	**.564***	.492	.418	6	/19
Polish ($\mathcal{S}$, L)	145 / 49 / 47	**.597**	.584	.594	**.555**	.542	.532	9	/19
Italian ($\mathcal{R}$, L)	227 / 76 / 61	**.584***	.502	.486	**.469****	.371	.372	5	/19
English ($\mathcal{G}$, L)	433 / 83 / 54	**.535***	.469*	.350	**.482***	.409*	.274	5	/23
French ($\mathcal{R}$, L)	158 / 53 / 50	**.469***	.463*	.329	**.429***	.419*	.276	9	/19
Russian ($\mathcal{S}$)	143 / 48 / 72	.409*	**.421***	.230	**.367****	.258	.218	5	/18
Spanish ($\mathcal{R}$, L)	– / – / 30	**.571****	.418*	.120	**.455****	.305*	.095	**1**	/17
Greek	– / – / 64	**.516***	.427	.345	**.410***	.338*	.057	7	/16
Georgian	– / – / 29	.400*	**.404***	.260	.291	**.384***	.251	9	/16
Summary	1238 /354 /457	**.523****	.471*	.356	**.447****	.391*	.277	$6.\overline{2}$	$/18.\overline{4}$

[C3] Reiter-Haas, M.*, Ertl, A.*, Innerebner, K., & Lex, E. (2023). mCPT at SemEval-2023 Task 3: Multilingual Label-Aware Contrastive Pre-Training of Transformers for Fewand Zero-shot Framing Detection. In Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023), pages 941–949, Toronto, Canada. Association for Computational Linguistics

RQ1: Framing Detection



RQ1c

[C2] Reiter-Haas, M. (2023). Exploration of Framing Biases in Polarized Online Content Consumption. In Companion Proceedings of the ACM Web Conference 2023, 560–564. Presented at the Austin, TX, USA.

RQ1c: Framing of COVID-19

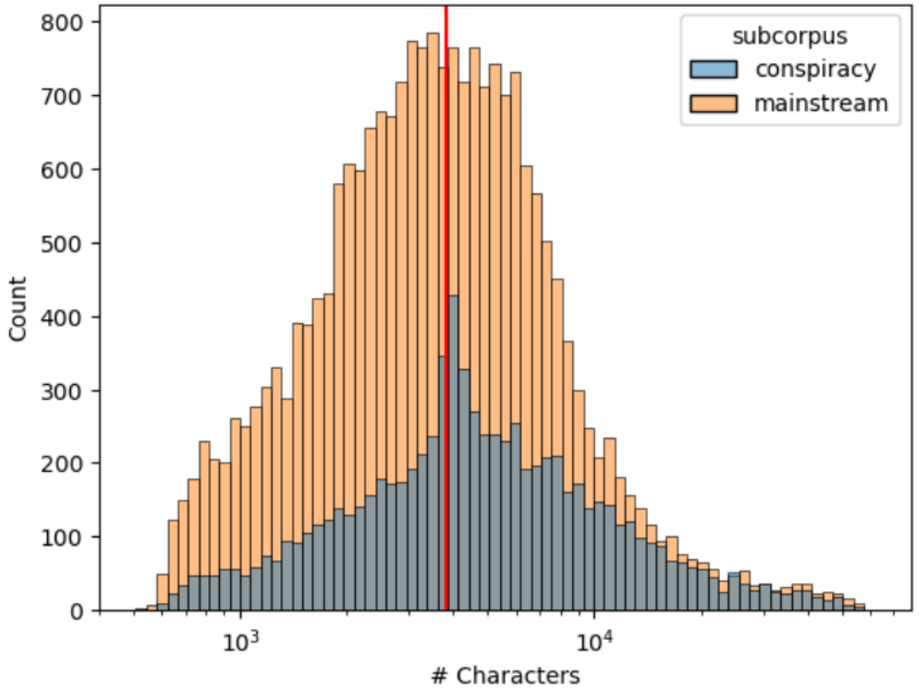
LOCO: mainstream vs conspiracy

Collected from news websites

Health (COVID-19, disease, pharma)

Semantic frames of 33,648 documents

Convert to graphs and mine differences

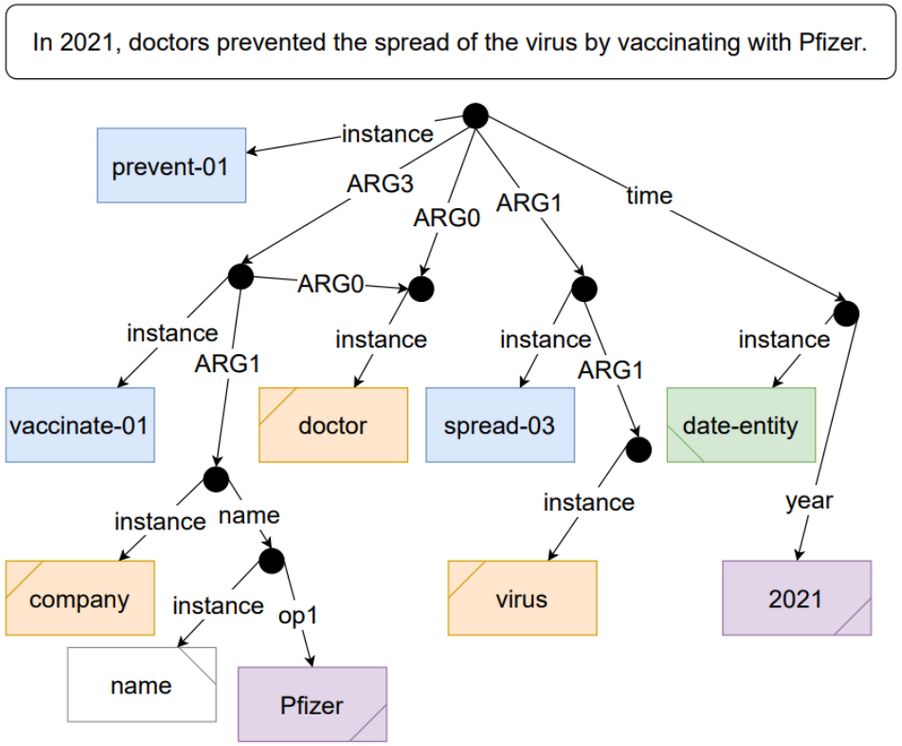


Miani, A., Hills, T., & Bangerter, A. (2021). LOCO: The 88-million-word language of conspiracy corpus. Behavior research methods, 1-24.

RQ1c: Abstract Meaning Representations (AMR)

Methodology:

- AMR parsing (BART model)
- Mining narrative structure
  - Characters
  - Plot
  - Setting
  - Moral of the Story
- Comparison of narrative information



Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., ... & Schneider, N. (2013, August). Abstract meaning representation for sembanking. In *Proceedings of the 7th linguistic annotation workshop and interoperability with discourse* (pp. 178-186).

RQ1c: Differences in Frame Structure

COVID-19 substructures:

Conspiracy ←→ Mainstream

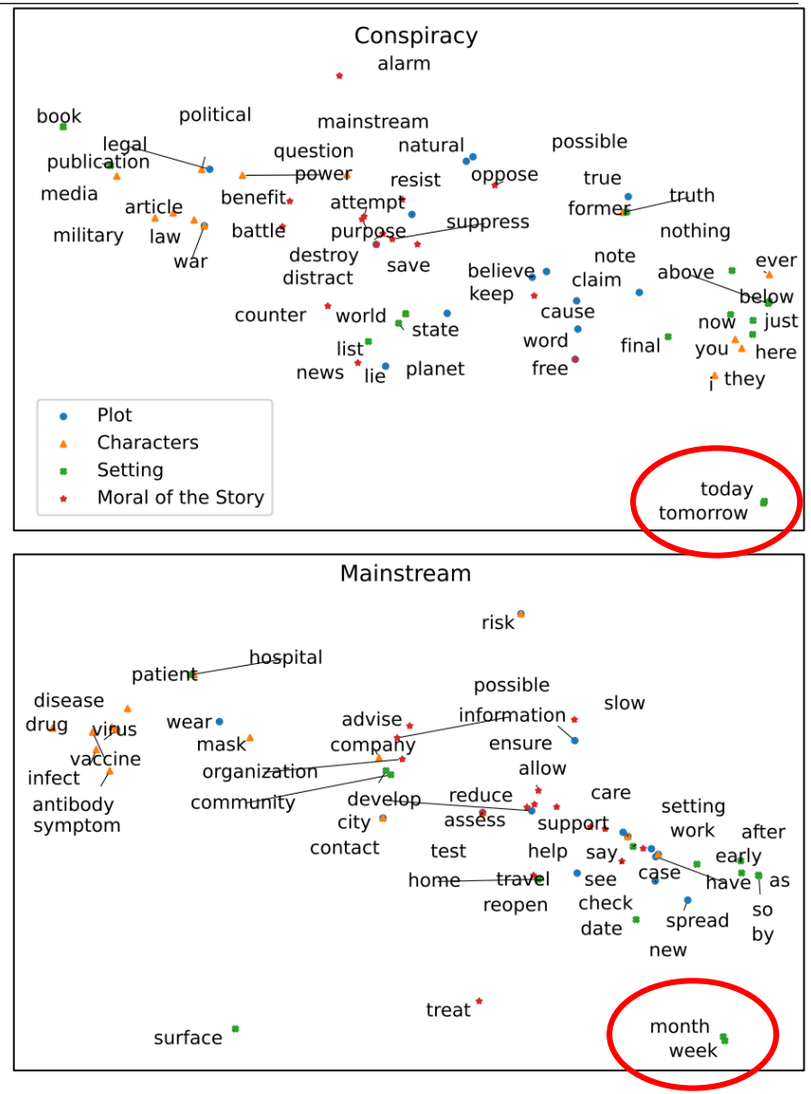
prevent-01 violence ←→ *prevent-01* infect-01

vaccine *spread-03* ←→ person *spread-03* virus

military *vaccinate-01* ←→ *vaccinate-01* person

Notable themes:

- Urgency and immediacy
- Science vs. belief orientation



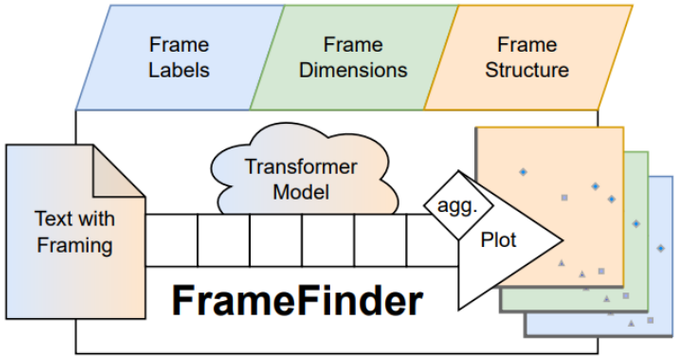
Embeddings space overrepresented of COVID-19 narrative elements

[C5] Reiter-Haas, M., Klösch, B., Hadler, M., & Lex, E. (2024). Framing Analysis of Health-Related Narratives: Conspiracy versus Mainstream Media. arXiv preprint arXiv:2401.10030.

RQ1: Exploration-Validation Trade-Off

	Labels (a)	Dimensions (b)	Structure (c)
Training Data	few* **samples**	pole information	none
Mode	supervised*	unsupervised	**discovery**
Extraction	scalar (per label)	n-Dimensional	irregular
Exploration	low	medium	high

* zero-shot also possible

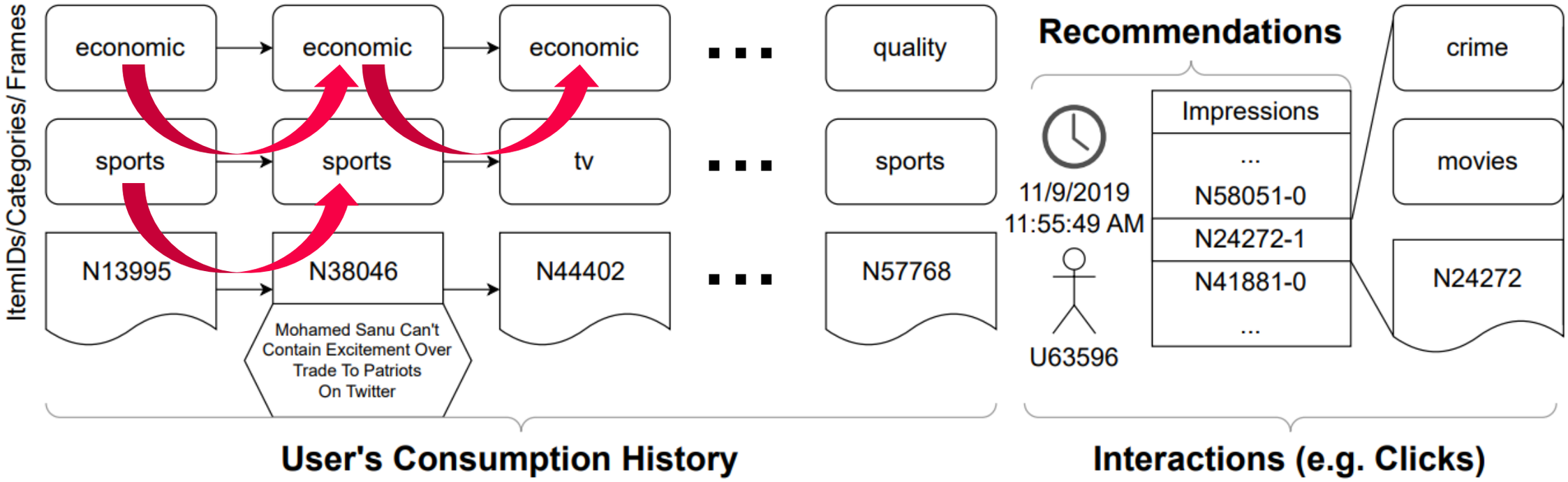


https://huggingface.co/spaces/Iseratho/frame-finder

[C6] Reiter-Haas, M., Klösch, B., Hadler, M., & Lex, E. (2024). FrameFinder: Explorative Multi-Perspective Framing Extraction from News Headlines. In Proceedings of the 2024 Conference of Human Information Interaction and Retrieval (CHIIR '24), 381-385.

RQ2: Information Behavior

MIND Dataset - Information System



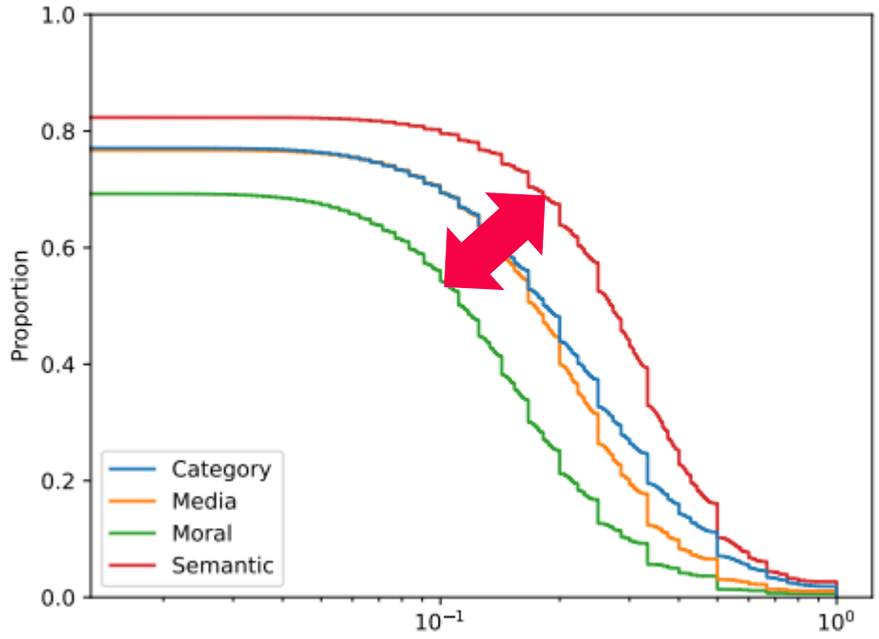
Wu, F., Qiao, Y., Chen, J. H., Wu, C., Qi, T., Lian, J., ... & Zhou, M. (2020, July). Mind: A large-scale dataset for news recommendation. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 3597-3606).

RQ2: Frame Consumption

Analyze Sequences:

- Repeat Consumption
- Viewpoint Diversity

Cat	IS
↑	↓
↓	↑



ECDF of the Direct Repetition Ratio (DDR)

1. Frame-specific consumption behavior
2. **Cat**egories are vital
3. Information systems (**IS**) positive effect due to choice

[C7] Reiter-Haas, M. & Lex, E. (2024). The Framing Loop: Do Users Repeatedly Read Similar Framed News Online? In Joint Proceedings of the ACM IUI Workshops 2024, March 18-21, 2024, Greenville, South Carolina, USA.

Talk Outline

1. Motivation of Framing
   - Research Questions
   - Overview of the Background
2. Publications
   - C1 – Polarization
   - C3 – Labels
   - C6 – Structure
   - C7 – Behavior
3. Conclusion
   - Future Research
   - PhD Summary

Summary of Contributions

Multi-perspective approach for framing detection — *RQ1 – Detection*

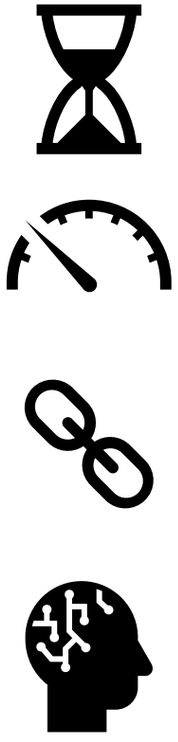
- Optimize embedding space by label overlap — *RQ1a – Labels*
- Convert to graph representations and reveal patterns — *RQ1c – Structure*

→ Exploration-Validation Trade-Off (**FrameFinder**)

Frame-specific influence on information behavior — *RQ2 – Behavior*

Future Work

- Temporal aspects
- Framing bias mitigation
- Unification
- The role of LLMs



Example for Narrative Framing:

Temporal evolution
in competing narratives
(from their narrative structure)
→ debias with LLMs

[C8] Reiter-Haas, M., Klösch, B., Hadler, M., & Lex, E. (2024). Computational Narrative Framing: Towards Identifying Frames through Contrasting the Evolution of Narration. In Proceedings of the Text2Story'24 Workshop, Glasgow (Scotland), 24-March-2024.

PhD Journey and Achievements

Research Output:
20+ Publications,
100+ Citations,
h-index of 6,
Open Source

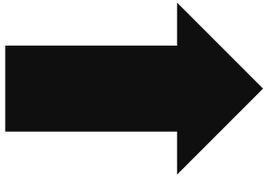
Reviewing Services:
80+ papers @
20+ CFPs
Best Reviewers @
TheWebConf

Teaching:
5 courses/events,
200+ students,
2 thesis co-supervision,
Community contributions

Events Participation:
10+ of which
4 physical scientific,
2 session chairs,
Erasmus+

Education:
25+ courses
(curricular, extra-curricular, teaching, summer schools)

Misc:
1st @ SemEval Spanish
Grants for Initial Funding and Travel
Outreach (e.g., press)



PostDoc: The Polarization Lab @ Duke University
Research on Polarization concerning the Framing on Social Media with Field Studies using LLMs

Thank you for your attention!

https://iseratho.github.io/thesis

Reiter-Haas, M. (2024). Computational Framing Analysis for Polarized Topics Online.

Additional References

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. science, 211(4481), 453-458.

Sullivan, K. (2023). Three levels of framing. Wiley Interdisciplinary Reviews: Cognitive Science, 14(5), e1651.