

# Teach Me How to Argue: A Survey on NLP Feedback Systems in Argumentation

**Camélia Guerraoui**, Paul Reisert, Naoya Inoue, Farjana Sultana Mim, Keshav Singh,  
Jungmin Choi, Irfan Robbani, Shoichi Naito, Wenzhi Wang and Kentaro Inui.

Tohoku University, RIKEN, INSA Lyon  
camelia.guerraoui@gmail.com

10<sup>th</sup> Workshop on Argumentation Mining - 7<sup>th</sup> December 2023

TOHOKU  
NLP LAB



# | Introduction

# Background: Necessity of critical thinking skills

- Digital era =  of information → Harder to evaluate its quality

→ Need to be able to *navigate* through this information  
= objectively analyze information and draw a rational conclusion  
= develop **critical thinking skills**

HOW?

- Strong argumentation → Strong critical thinking skills

(Pithers and Soden, 2000; Behar-Horenstein and Niu, 2011)

# Computational argumentation

- **Computational argumentation** assists users in improving their arguments
- Various applications:
  - Mining arguments (Al-Khatib et al., 2016)
  - Assessing arguments' quality (El Baff et al., 2018)
  - Reconstructing implicit assumptions in arguments (Habernal et al., 2018)
  - etc.
- Assist students' learning & reduce teachers' workload (Twardy, 2004; Wambsganß et al., 2021)
  - But **cannot still deeply explain and visualize** how an argument can be improved!

# Gap between current and desired argumentative systems

- **Aim:** Exploration of current argumentative feedback systems for improving critical thinking skills
- **Contributions:**
  - Combine features of feedback systems into **four distinct dimensions**
  - Survey and categorize **108 papers** into these dimensions
  - Discuss remaining **challenges** and potential ways to overcome them
  - Creation of **a website** to easily find our references

[https://kmilia.github.io/teach\\_me\\_how\\_to\\_argue/](https://kmilia.github.io/teach_me_how_to_argue/)



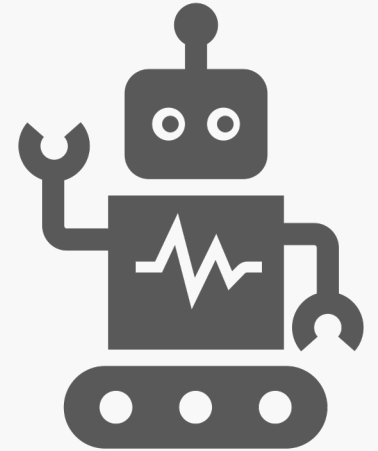
## | Dimensions of the survey


## Richness of explanations = **What?**



 cuisine is not healthy.  
Indeed  are not healthy.

Your argument has an invalid  
generalization.





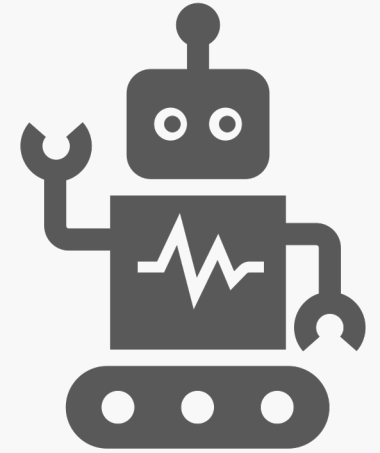
- Shallow feedback
- Not enough details... 





# Richness of explanations = Why?



 cuisine is not healthy.  
Indeed  are not healthy.

Your argument has an invalid  
generalization.  
 **cuisine is not just** 






- Easier to understand 
- Need for implicit knowledge:  
Cobb salad = healthy  salad
- Use of LLMs (Examples: Dalvi Mishra et al., 2023)
  -  Find implicit components
  -  Find the correct logical reasoning
  - Need for new techniques  
(E.g: chain-of-thought prompting (Wei et al., 2022))

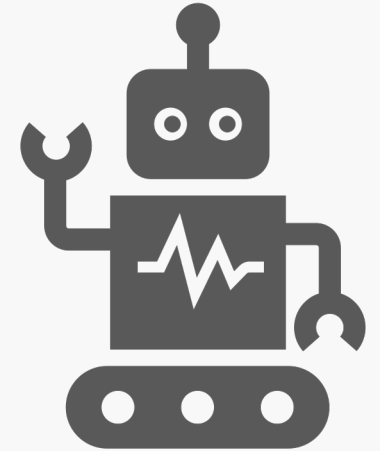


# Visualization of explanations = How?



 cuisine is not healthy.  
*Indeed*  *are not healthy.* 

 Your argument has an *invalid generalization.*  
 cuisine is not just 





- Multiple views (Example: Wambsganss et al., 2020)
  - Diagrams → Argumentative structure
  - Low accuracy

- Dialogue Systems (Example: Rach et al., 2020)
  - Gaining attraction
  - Challenge: user-friendliness

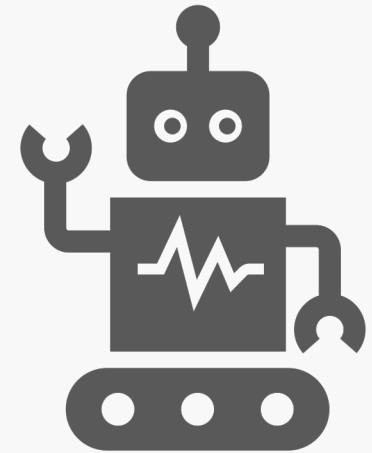
# Interaction = **Who** is interacting **with whom**?




 cuisine is not healthy.  
Indeed  are not healthy.

Your argument has an invalid  
generalization.  
 cuisine is not just 

**What do you mean?**



- **Between users** (Example: Lugini et al, 2020)
  - Users debating/helping each other
  - Challenge: Need a real-class setting

- **With a chatbot** (Example: Wambsganss et al., 2021)
  - Based on state-of-the-art LLMs
  -  Generate human-like responses

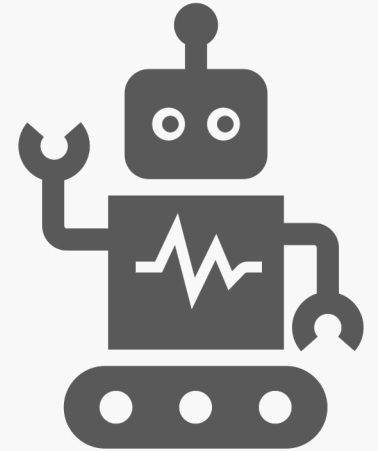
# Personalization of the explanations = To whom?



 cuisine is not healthy.  
Indeed  are not healthy.

Careful, you are simplifying!

 cuisine is not just 



- Discretization of users' levels

(Example: Wachsmuth and Alshomary, 2022)

- No user's background

- Personalization by users

(Example: Putra et al., 2020)

- Not very user-friendly

- Use of chatbots (Example: Ismail et al., 2023)

- Tailor their responses to learners
- Possible future direction

| **General open issues**

# General challenges identified while surveying

- Evaluating different systems
  - Relies on human → hard to reproduce and costly
  - Lack of direct comparisons between similar systems (Heuer and Buschek, 2021)
  - Need to promote open-source projects and the research of standard guidelines
- Ethics
  - Challenges: Privacy violations, discrimination, bias, etc (Hovy et al., 2017; Trust et al., 2023)
  - Need for an ethics-by-design approach (Leidner and Plachouras, 2017)
- ... More to be discussed during the Poster session

| **Conclusion**

# Summary

- Overview of current systems used to teach argumentation based on:

1. **What** is an error?  
**Why** is it an error? } → **Richness**
2. **How** is the error visualized? → **Visualization**
3. **Who** is interacting **with whom**? → **Interaction**
4. **To whom** is given the feedback? → **Personalization**

- Identified several open issues: Evaluation Methods, Reasoning...

# Towards an end-to-end system for the user

The screenshot shows a web browser window with a 'New Tab' and a search bar. The page title is 'Previous debates Start new debate' and the user is 'Flavie Murea'. The main content is a debate interface for the theme 'Homework'. On the left, there is a text area for the 'Opponent' with a 'Write a counter-argument attacking this argument.' prompt. The main text area contains a draft argument about homework. On the right, there is a chat interface with a robot icon representing 'Deb8'. The chat shows a message from Deb8: 'I'm Deb8, your new assistant during your debate. What do you want to do?'. Below this, there are buttons for 'Ask a question', 'Correct my draft', and 'Me'. Another message from Deb8 says: 'Your draft has been corrected. Suggestions are highlighted in your text. For more information, please click on the hints. What do you want to do?'. Below this, there are buttons for 'Ask a question', 'Correct my draft', and 'Me'. An error message is also visible: 'Error type: Incomplete Premise. The opposing side might argue that the goal of university admission can be achieved even without homework. Save Dismiss'.

- **Goal:** prototype a system to improve students' critical thinking

- **Measure:**

$$f(\text{feedback} + \text{interface}) =$$

Which effects on learning critical thinking?

Preliminary sketch of an end-to-end system to learn argumentation



# | Appendix

# Argumentation : a definition

**Argumentation** = Field of elaborating and presenting arguments to engage in debate, convince others, and eventually reach agreements



**Claim**

**Premise**

## 4 Pedagogical methods = **How** shall we teach?



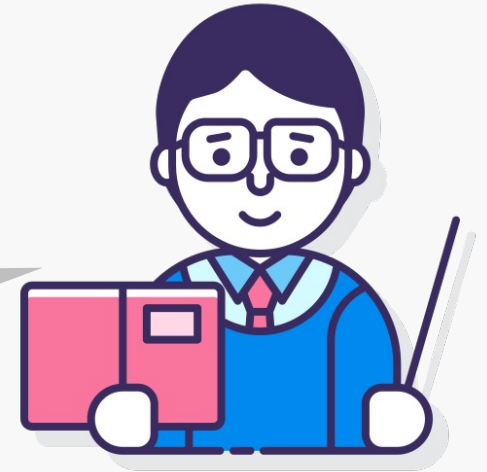
cuisine is not healthy.

Indeed



are not healthy.

Your argument has an invalid  
generalization.



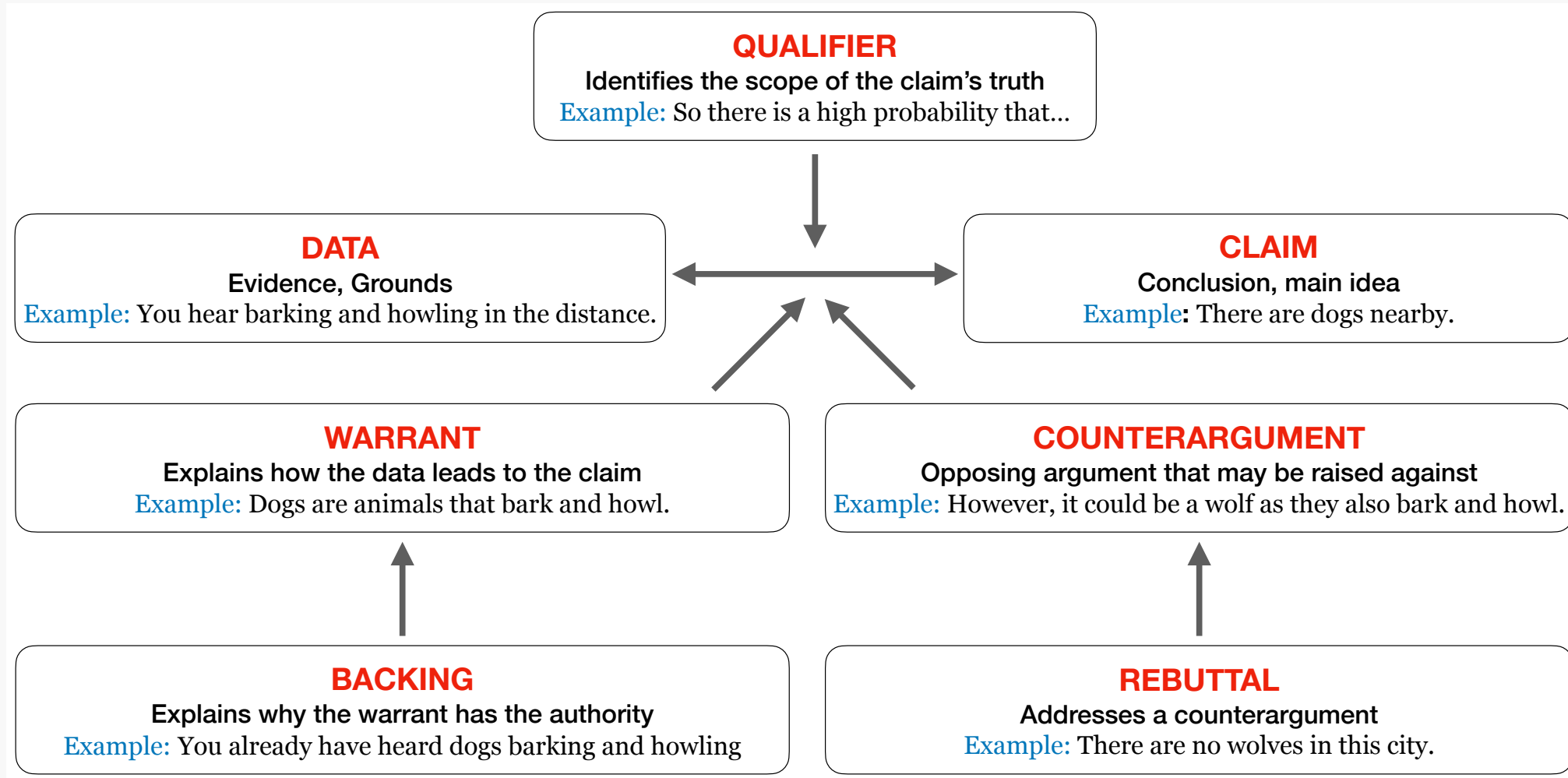
1. Toulmin Model

2. Rhetorical structure theory

3. Collaborative argumentation

4. Socratic questioning

# Toulmin Model



# Survey method

- Deeply read 150 papers collected in an Excel file
- Map each paper to at least one category
- Whole Survey: 108 cited papers



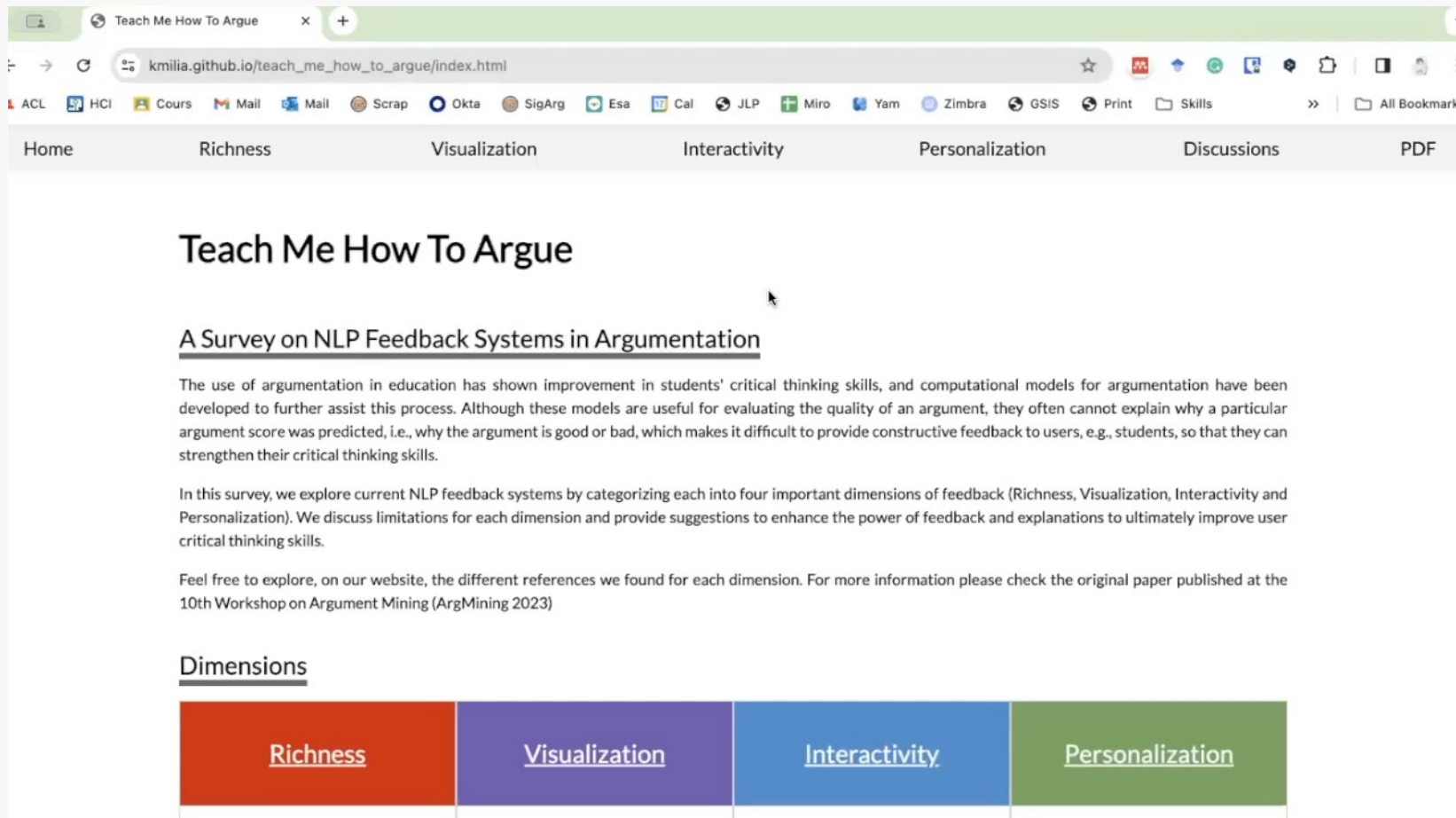
Papers (with link)	Conference	Year	What/Why	How	To whom	Who	Who	Who
			Richness	Visualization	Personalization	Interactivity	Trigger	Role / Responsibility
<a href="#">A PhD Student's Perspective on Research in NLP in the Era of Very Large Language Models</a>		2023	Gaps and potential directions for further research in NLP and specifically with LLMs					
<a href="#">Annotating Arguments: The NOMAD Collaborative Annotation Tool</a>	LREC	2014						Pioneer in collaborative tools for improving arguments
<a href="#">VISAR: A Human-AI Argumentative Writing Assistant with Visual Programming and Rapid Draft Prototyping</a>	-	2023		Graph Component dynamically linked to a text editor, highlighted argumentative components				Future work: collaborative settings
<a href="#">ArgueTutor: An Adaptive Dialog-Based Learning System for Argumentation Skills</a>	CHI	2021		Chat No Clear history		Possibility of interacting with the system	The user asks for the feedback	
<a href="#">AI: An adaptive learning support system for argumentation skills.</a>	CHI	2020		Graphs + Scores + History + Visualization of Claim and premises	5 levels of feedback (Novice, Advanced, Competent, Proficient, Expert)			
<a href="#">"mama always had a way of explaining things so I could understand": A dialogue corpus for learning to construct explanations.</a>	COLING	2022			5 levels of explanations	Studies to see how humans explain in dialogical settings		

Dimensions	Pedagogy <b>How</b>	Richness <b>What/Why</b>	Visualization <b>How</b>	Interactivity <b>Who</b>	Personalization <b>To Whom</b>	Total
References	19	32	13	10	11	85

Table 1: Cumulative count per dimension of papers referenced in our survey.

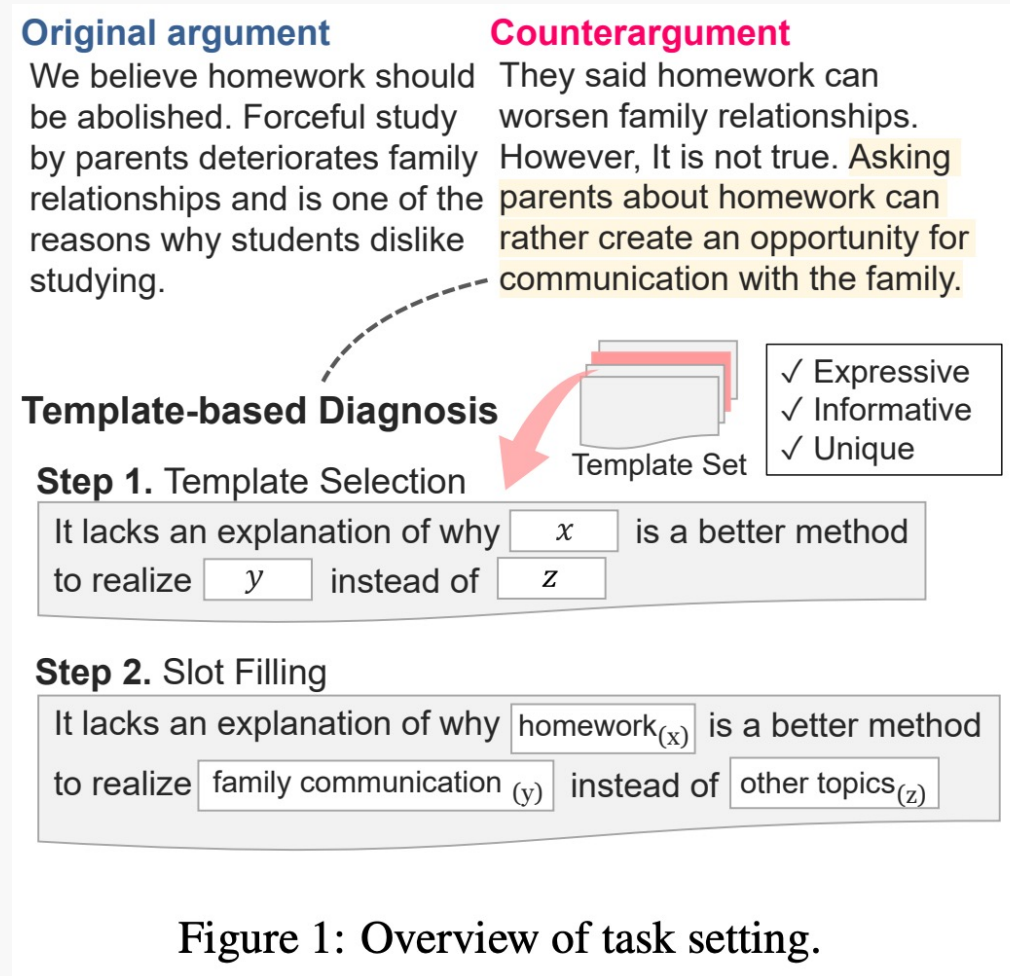
# Easy access to our references: Creation of a website

All references can be found at [https://kmilia.github.io/teach\\_me\\_how\\_to\\_argue/](https://kmilia.github.io/teach_me_how_to_argue/)



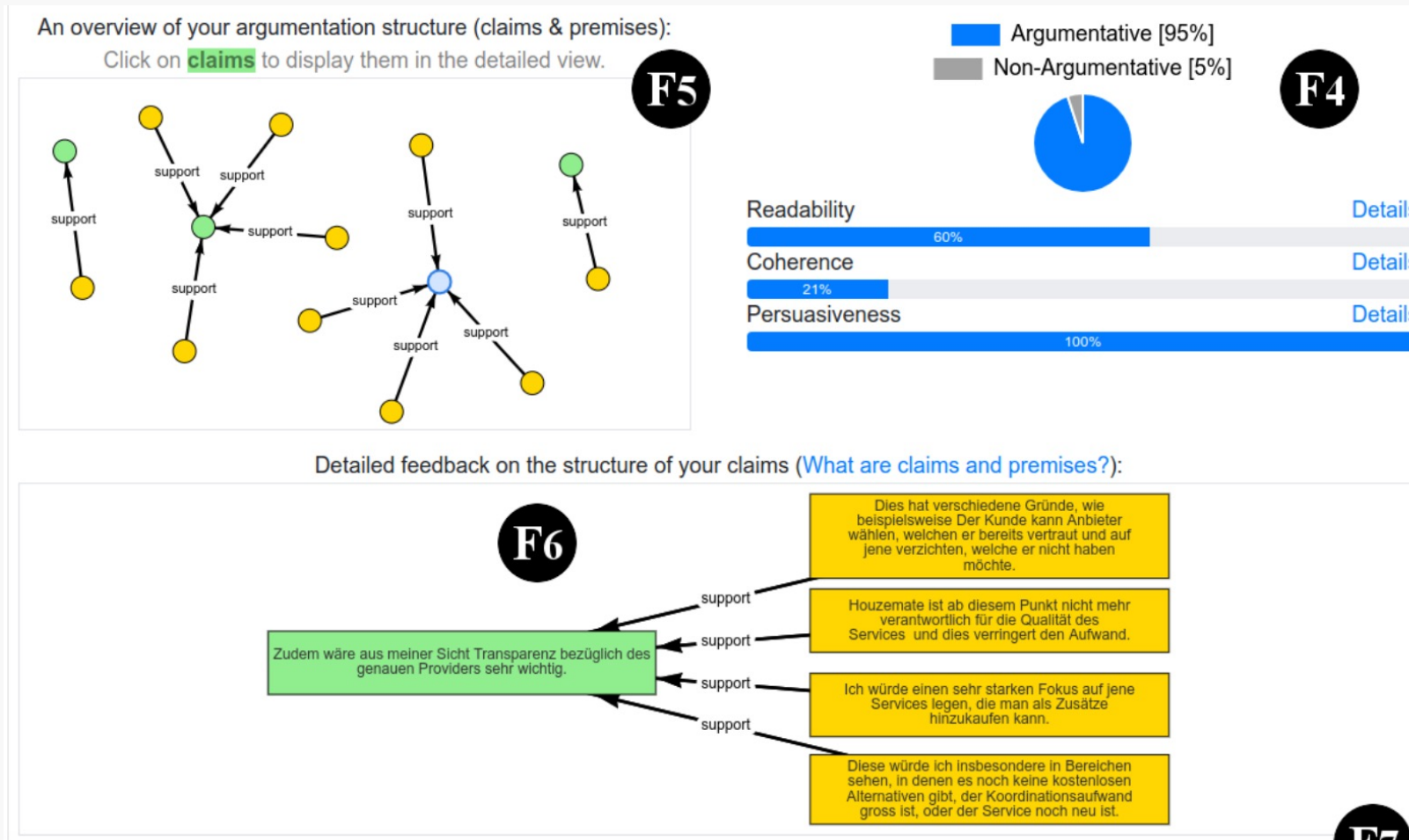
# Richness of Explanations = What/Why? - Example

TYPIC: A Corpus of Template-Based Diagnostic Comments on Argumentation,  
S. Naito et al. EMNLP 2022



# Visualization of Explanations = How? - Example

AL: An Adaptive Learning Support System for Argumentation Skills,  
T. Wambsganss et al. CHI 2020





# Interaction = Who is interacting with whom? - Example

ArgueTutor: An Adaptive Dialog-Based Learning System for Argumentation Skills,  
T. Wambsganss et al. CHI 2021

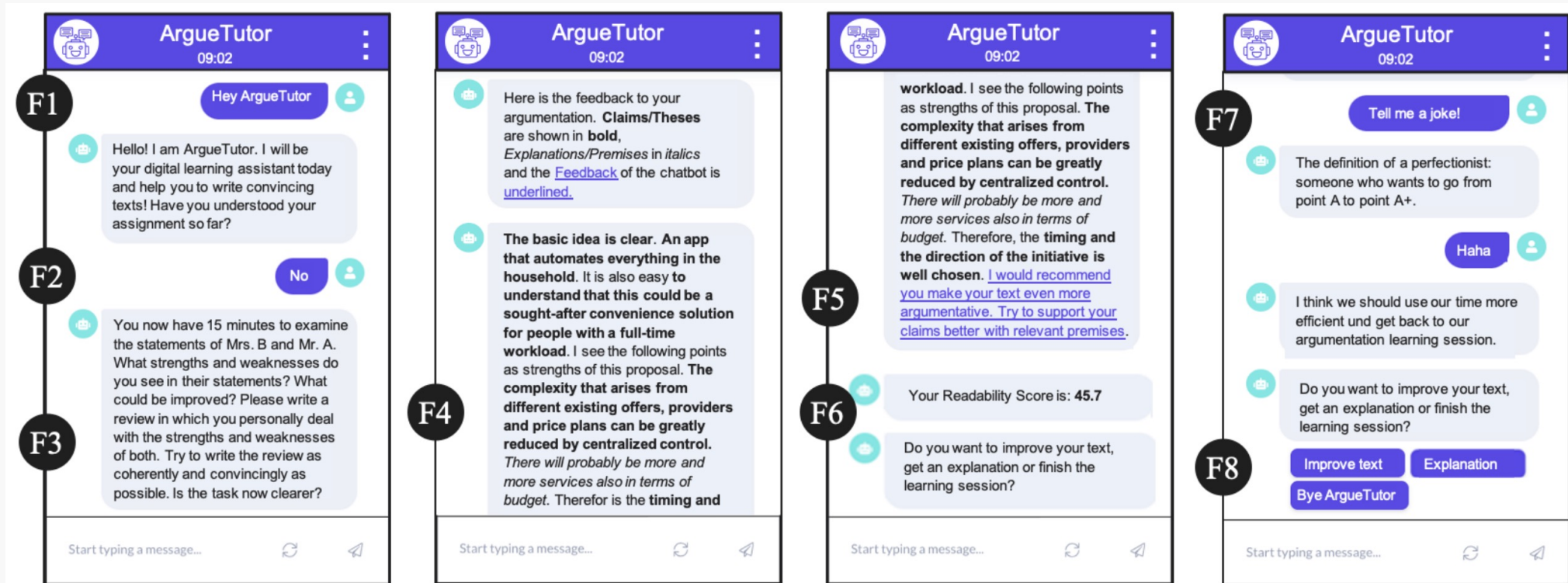


Figure 1: Screenshot of our adaptive dialog-based learning system: a user conducts a certain writing exercise and receives adaptive tutoring and feedback on the argumentation quality of her text

# Personalization of the Explanations = To Whom? - Example

“Mama Always Had a Way of Explaining Things So I Could Understand”: A Dialogue Corpus for Learning to Construct Explanations, H. Wachsmuth et al. COLING 2022

**Explaining dialogue on the main topic “blockchain”**

01 Do you know what we're gonna talk about today? It's called blockchain.

02 What's blockchain?

03 That's a really good question. It's actually a way that we can trade. Do you know what trade is?

04 Mmm-hmm, it's when you take turns doing something. It's when you give up most of what you want, right?

05 When you give up most of what you want? Well, sometimes that definitely happens for sure. What if I told you that this is the kind of technology that I work on that means you could trade with any kid all over the world?

06 Really?

07 Yeah.

08 If I could trade with any kid, I would trade, well, I would trade something I don't like so much.

09 That's probably a good idea, maybe somebody else likes it more than you do. So normally, when people trade, they have to go to the store, or they have to know the person so they can get what they asked for. With blockchain, you can make that exact same trade, but you don't need the store, and you don't even necessarily need to know the other person.

10 Really?

11 Really.

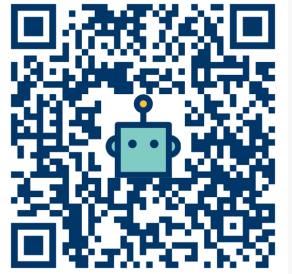
**Explainer (expert)** (child) **Explainee**

Figure 1: A short explaining dialogue from the video series *5 Levels*, included in the corpus presented in Section 3. Here, an expert explains blockchain to a child.

# Related surveys in argumentation

- **Automated Writing Evaluation** (Ke and Ng, 2019; Beigman Klebanov and Madnani, 2020; Wang et al., 2022)
  - Mainly about essay scoring → Shallow feedback
  - Not deeply covering the different argumentative feedback
- **Argumentation Mining** (Habernal and Gurevych, 2016; Lawrence and Reed, 2020)
  - Point of view of the coder
  - Not the point of view of the user (not feedback-oriented)
- **Explainability** (Vassiliades et al. 2021; Čyras et al. 2021)
  - Show the potential of argumentation to explain a model output
  - But don't explain how to explain an argument

# References



All references can be found at [https://kmilia.github.io/teach\\_me\\_how\\_to\\_argue/](https://kmilia.github.io/teach_me_how_to_argue/)

- *Teaching Critical Thinking Skills In Higher Education: A Review Of The Literature*, Behar-Horenstein et al, 2011
- *Argument maps improve critical thinking*. Twardy. Teaching Philosophy, Vol. 27, , 06 2004
- *Computer-supported argumentation: A review of the state of the art*. Scheuer et al., 2010
- *LPAttack: A Feasible Annotation Scheme for Capturing Logic Pattern of Attacks in Arguments*, Mim et al. EMNLP 2022
- *TYPIC: A Corpus of Template-Based Diagnostic Comments on Argumentation*, Naito et al. EMNLP 2022
- *AL: An Adaptive Learning Support System for Argumentation Skills*, Wambsganss et al. CHI 2020
- *Arguetutor: An adaptive dialog-based learning system for argumentation skills*. Wambsganß, et al., 2021
- *“Mama Always Had a Way of Explaining Things So I Could Understand”*: A Dialogue Corpus for Learning to Construct Explanations, Wachsmuth et al. COLING 2022