



<https://tinyurl.com/bambergtranslate>



Natural Language Processing Research in the Era of AI

Can Computers Read Between the Lines?

Inaugural Lecture, November 18, 2025

Roman Klinger

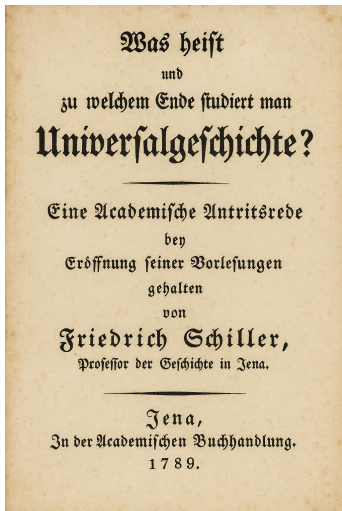
roman.klinger@uni-bamberg.de

 [romanklinger.de](https://twitter.com/romanklinger)  [romanklinger](https://www.linkedin.com/in/romanklinger)

<https://www.bamberg.de/nlproc/>

<https://www.romanklinger.de/>





“What Is, and to What End do
We Study, Universal History?”

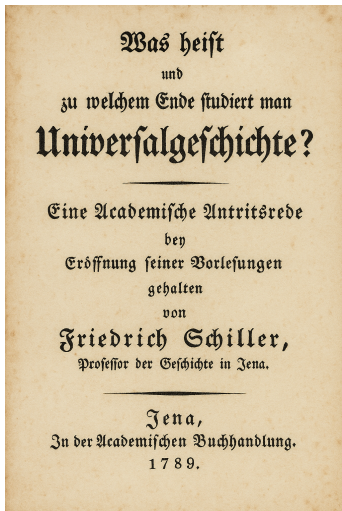
Was heist
und
zu welchem Ende studiert man
Universalgeschichte?

Eine Academische Antrittsrede
bey
Eröffnung seiner Vorlesungen
gehalten
von
Friedrich Schiller,
Professor der Geschichte in Jena.

Jena,
In der Academischen Buchhandlung.
1789.

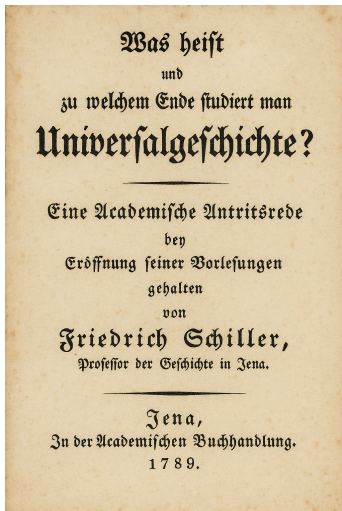
“What Is, and to What End do
We Study, Universal History?”

- Schiller contrasts two types of students:



“What Is, and to What End do We Study, Universal History?”

- Schiller contrasts two types of students:
 - **Brotgelehrter** (“bread scholar”) who cares about practical advantage and career.



“What Is, and to What End do We Study, Universal History?”

- Schiller contrasts two types of students:
 - **Brotgelehrter** (“bread scholar”) who cares about practical advantage and career.
 - **Philosophischer Kopf** (“philosophical mind”) who seeks knowledge and human betterment.

What do we do today?



University of Bamberg



Natural Language Processing Research in the Era of AI

Can Computers Read Between the Lines?

Inaugural Lecture, November 18, 2025

Roman Klinger

roman.klinger@uni-bamberg.de

 romanklinger.de  [romanklinger](https://www.linkedin.com/in/romanklinger)
<https://www.bamberg.de/nlproc/>

What do we do today?



University of Bamberg



Natural Language Processing Research in the Era of AI

Can Computers Read Between the Lines?

Inaugural Lecture, November 18, 2025

Roman Klinger
roman.klinger@uni-bamberg.de

 romanklinger.de  [romanklinger](https://www.linkedin.com/in/romanklinger)
<https://www.bamberg.de/nlproc/>

- What is, and why do we study, Natural Language Processing?

What do we do today?



University of Bamberg



Natural Language Processing Research in the Era of AI

Can Computers Read Between the Lines?

Inaugural Lecture, November 18, 2025

Roman Klinger
roman.klinger@uni-bamberg.de

 romanklinger.de  [romanklinger](https://www.linkedin.com/in/romanklinger)
<https://www.bamberg.de/nlproc/>

- What is, and why do we study, Natural Language Processing?
- How does this research field change in the era of "AI"?

Outline

1 Language Understanding

2 What is “AI”?

3 What we work on

4 Wrap Up

Outline

1 Language Understanding

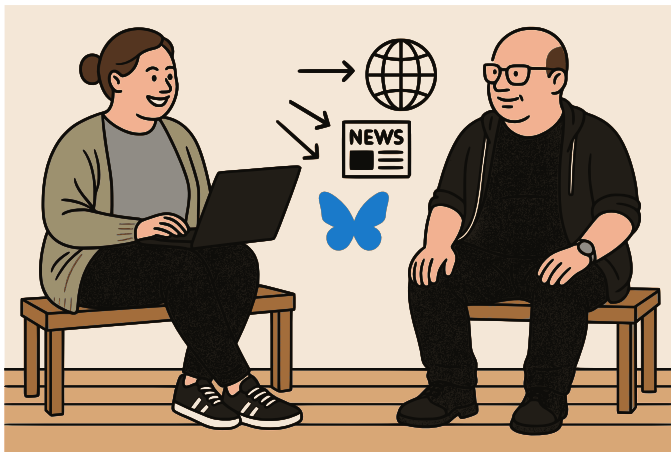
2 What is “AI”?

3 What we work on

4 Wrap Up

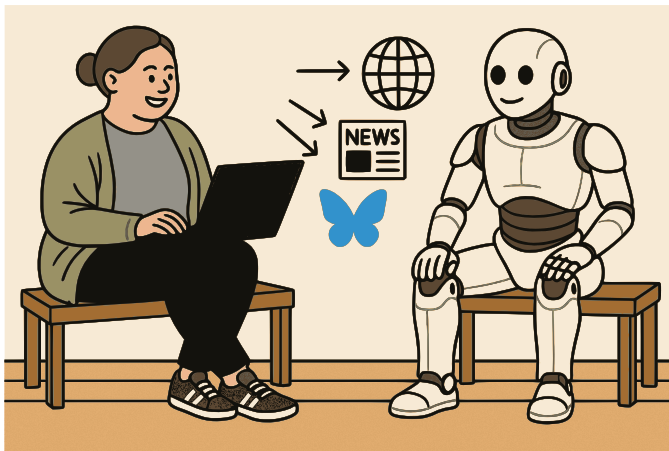


Natural Language Processing and Understanding



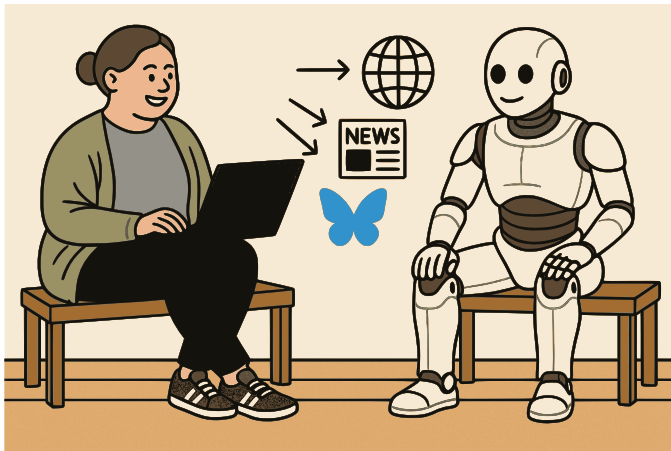


Natural Language Processing and Understanding





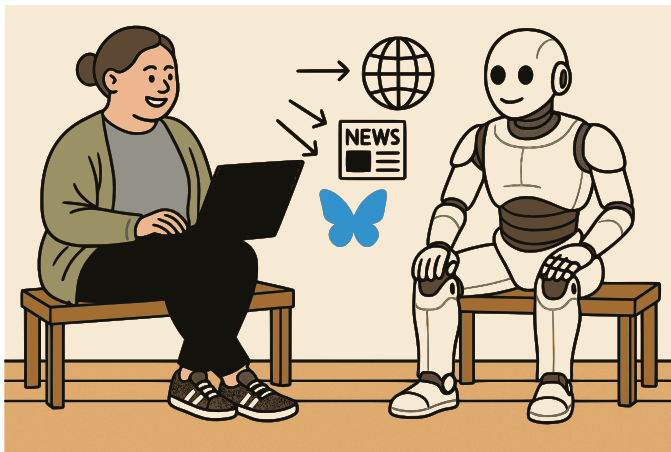
Natural Language Processing and Understanding



- We study how machines can understand human language



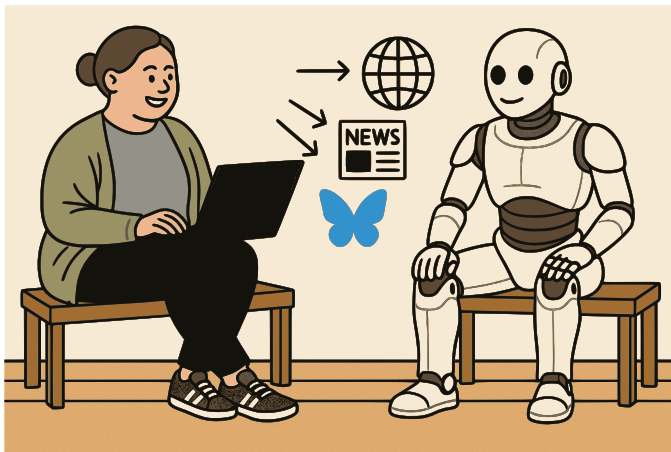
Natural Language Processing and Understanding



- We study how machines can understand human language
- We focus on written text



Natural Language Processing and Understanding



- We study how machines can **understand** human language
- We focus on written text

Understanding...



Understanding...



- What does it mean to “understand”?

Understanding...



- What does it mean to “understand”?
- Cambridge Dictionary: “to know the meaning of something that someone says”

Understanding...



- What does it mean to “understand”?
- Cambridge Dictionary: “to know the meaning of something that someone says”
- How can we make computers understand?

Understanding...



- What does it mean to “understand”?
- Cambridge Dictionary: “to know the meaning of something that someone says”
- How can we make computers understand?
- How can we measure if we are successful?

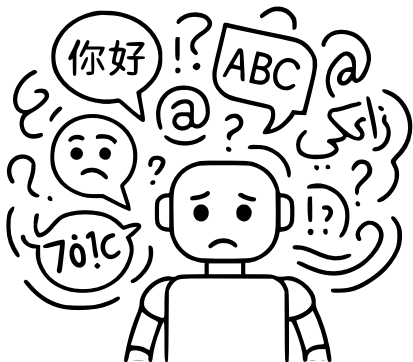
Understanding...



- What does it mean to “understand”?
- Cambridge Dictionary: “to know the meaning of something that someone says”
- How can we make computers understand?
- How can we measure if we are successful?
- How and what for can we make use of the meaning that has been understood?



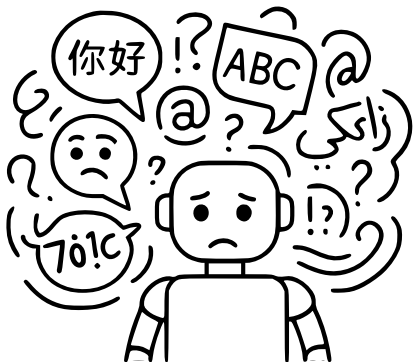
Natural Language Understanding



- Desideratum: A machine that **understands language as humans do?**



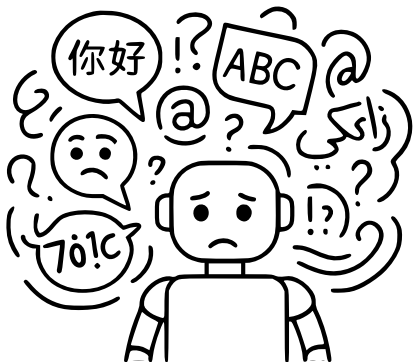
Natural Language Understanding



- Desideratum: A machine that **understands language as humans** do?
- How to study language in its entirety?
(**universal language understanding ability**)



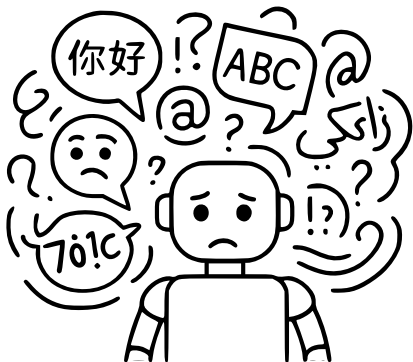
Natural Language Understanding



- Desideratum: A machine that **understands language as humans** do?
- How to study language in its entirety?
(**universal language understanding ability**)
- We study **particular phenomena**.



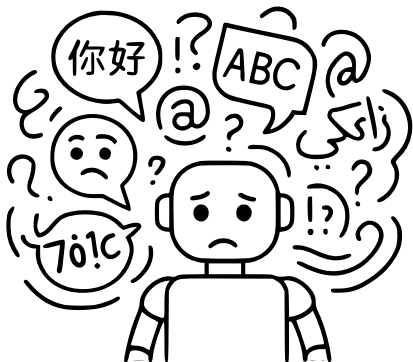
Natural Language Understanding



- Desideratum: A machine that **understands language as humans** do?
- How to study language in its entirety? (**universal language understanding ability**)
- We study **particular phenomena**.
- We define **concrete tasks to solve**.



Natural Language Understanding



- Desideratum: A machine that **understands language as humans** do?
 - How to study language in its entirety? (**universal language understanding ability**)
 - We study **particular phenomena**.
 - We define **concrete tasks to solve**.
- ⇒ Pragmatic approach to language understanding



World News > United States



Donald Trump mocks Bill Gates after billionaire's humiliating backflip on climate change

30 October 2025 - 04:46AM

sky news .com.au

Donald Trump wasted no time mocking Bill Gates after the billionaire admitted that climate change will not lead to humanity's demise in a memo released yesterday.



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names:



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment:



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment: **negative**



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment: **negative**
- Topic:



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment: **negative**
- Topic: **climate change**



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment: **negative**
- Topic: **climate change**
- Stances:



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment: **negative**
- Topic: **climate change**
- Stances: **Bill Gates** → opinion change on **climate change**.



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment: **negative**
- Topic: **climate change**
- Stances: **Bill Gates** → opinion change on **climate change**.
- Relation:



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment: **negative**
- Topic: **climate change**
- Stances: **Bill Gates** → opinion change on **climate change**.
- Relation: **Donald Trump** → **negative opinion** (**Bill Gates**)



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment: **negative**
- Topic: **climate change**
- Stances: **Bill Gates** → opinion change on **climate change**.
- Relation: **Donald Trump** → **negative opinion** (**Bill Gates**)
- ...



Tasks in Natural Language Understanding

“Donald Trump mocks Bill Gates after
billionaire’s humiliating backflip on climate change”

Sky News Australia, Oct 30, 2025

What information is in this sentence that’s worth understanding?

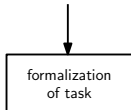
- Find entity names: **Donald Trump** ; **Bill Gates**
- Recognize the sentiment: **negative**
- Topic: **climate change**
- Stances: **Bill Gates** → opinion change on **climate change**.
- Relation: **Donald Trump** → **negative opinion** (**Bill Gates**)
- ...
- Aggregating information enables many use cases:
diverse news recommendation, social network analysis, opinion mining, ...

How to build such systems?



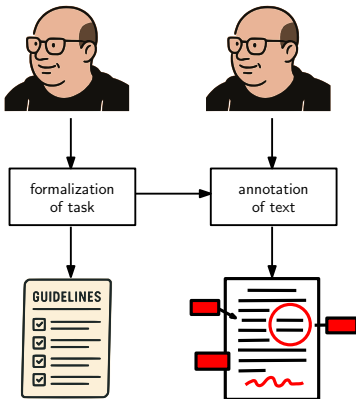


How to build such systems?



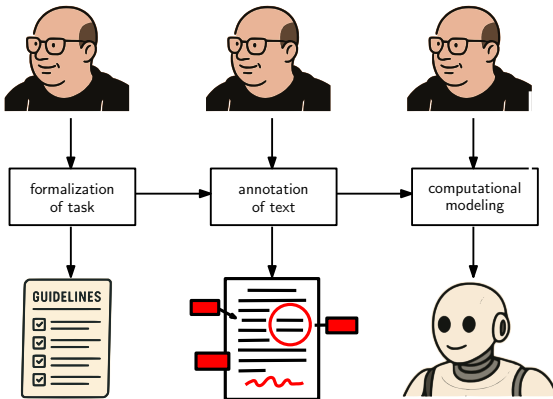


How to build such systems?



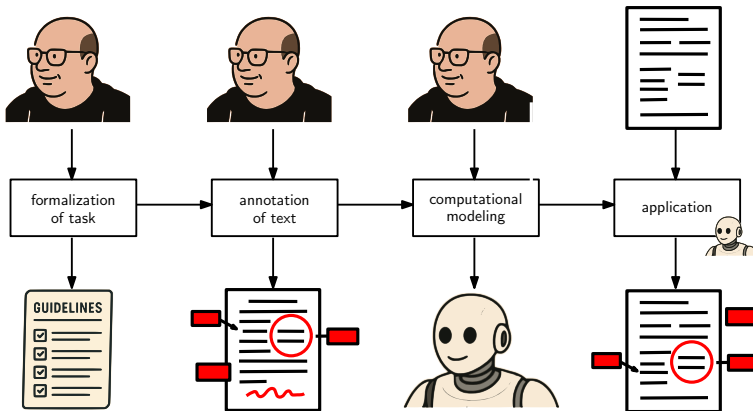


How to build such systems?



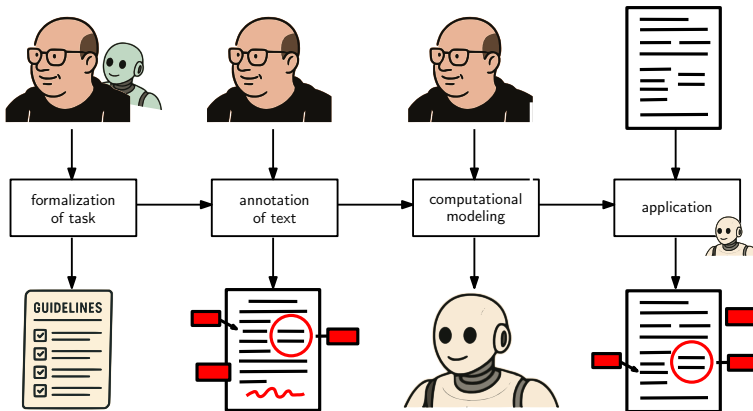


How to build such systems?



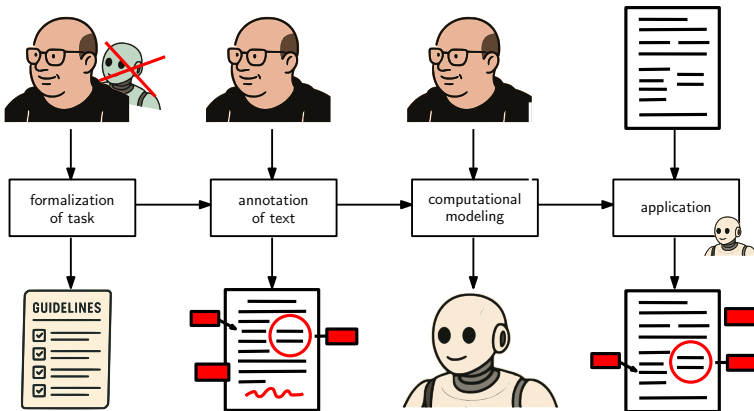


How to build such systems?



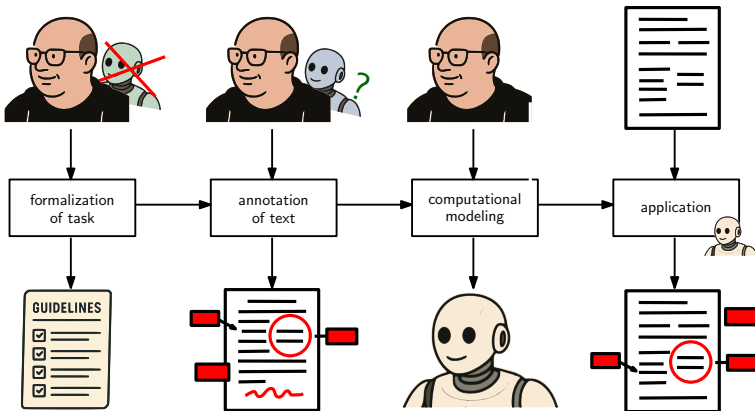


How to build such systems?



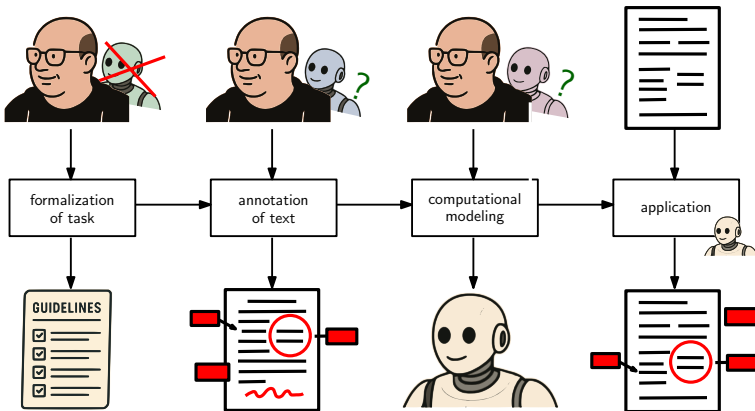


How to build such systems?



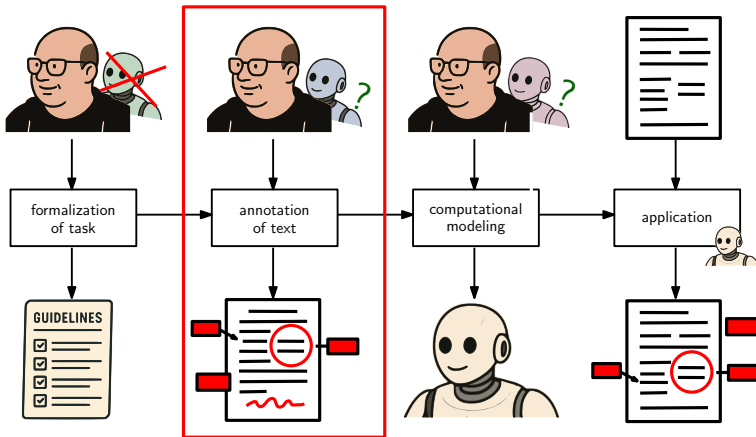


How to build such systems?





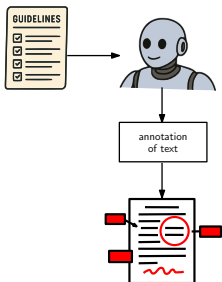
How to build such systems?





Do we still need the human?

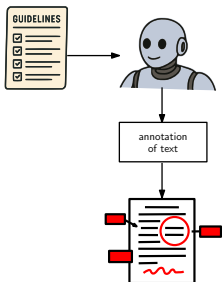
(except for the specification of the task)





Do we still need the human?

(except for the specification of the task)

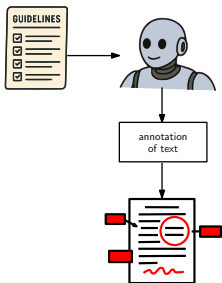


- If an “AI” can just solve the task by being instructed, we don’t need manually annotated data, and we don’t need to spend a lot of effort in model development.



Do we still need the human?

(except for the specification of the task)



- If an “AI” can just solve the task by being instructed, we don’t need manually annotated data, and we don’t need to spend a lot of effort in model development.
- In other words:
Does Bamberg need a Natural Language Processing professor?



**I DO NOT
NEED YOU!**

That is a reasonable question! We should reflect on our role.



That is a reasonable question! We should reflect on our role.



A lot of research goes into:

That is a reasonable question! We should reflect on our role.



A lot of research goes into:

1. Formalizing concepts



That is a reasonable question! We should reflect on our role.

A lot of research goes into:

1. Formalizing concepts
2. Collecting and annotating corpora for model development



That is a reasonable question! We should reflect on our role.

A lot of research goes into:

1. Formalizing concepts
2. Collecting and annotating corpora for model development
3. Advancing machine learning methods to better learn to reconstruct the annotations

That is a reasonable question! We should reflect on our role.



A lot of research goes into:

1. Formalizing concepts
2. Collecting and annotating corpora for model development
3. Advancing machine learning methods to better learn to reconstruct the annotations
4. Evaluation methods



That is a reasonable question! We should reflect on our role.

A lot of research goes into:

1. Formalizing concepts
2. Collecting and annotating corpora for model development
3. Advancing machine learning methods to better learn to reconstruct the annotations
4. Evaluation methods

If “AI” can do these steps sufficiently well, our work is essentially done.

Is that the case?

Outline

1 Language Understanding

2 What is “AI”?

3 What we work on

4 Wrap Up

Is our work done? What is “AI”?



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.





Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.
- Many agree that an important aspect is:
a system behaves in a way that appears to be intelligent.



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.
- Many agree that an important aspect is:
a system behaves in a way that appears to be intelligent.
- Methods:



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.
- Many agree that an important aspect is:
a system behaves in a way that appears to be intelligent.
- Methods:
 - Logical reasoning



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.
- Many agree that an important aspect is:
a system behaves in a way that appears to be intelligent.
- Methods:
 - Logical reasoning
 - Machine learning



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.
- Many agree that an important aspect is:
a system behaves in a way that appears to be intelligent.
- Methods:
 - Logical reasoning
 - Machine learning
 - Semantic language representations



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.
- Many agree that an important aspect is:
a system behaves in a way that appears to be intelligent.
- Methods:
 - Logical reasoning
 - Machine learning
 - Semantic language representations
 - Search and planning



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.
- Many agree that an important aspect is:
a system behaves in a way that appears to be intelligent.
- Methods:
 - Logical reasoning
 - Machine learning
 - Semantic language representations
 - Search and planning
 - Integration of world knowledge



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.
- Many agree that an important aspect is:
a system behaves in a way that appears to be intelligent.
- Methods:
 - Logical reasoning
 - Machine learning
 - Semantic language representations
 - Search and planning
 - Integration of world knowledge
 - ...



Is our work done? What is “AI”?

- To answer that, we need to understand what “Artificial Intelligence” means.
- The term encompasses many different topics.
- Many agree that an important aspect is:
a system behaves in a way that appears to be intelligent.
- Methods:
 - Logical reasoning
 - Machine learning
 - Semantic language representations
 - Search and planning
 - Integration of world knowledge
 - ...
- ⇒ The term “Artificial Intelligence” does NOT refer to one technology, but to many.

A common interpretation of “AI” in the moment: Language Models





A common interpretation of “AI” in the moment: Language Models

What word fits?



A common interpretation of “AI” in the moment: Language Models

What word fits?

1. “I came here by _____”



A common interpretation of “AI” in the moment: Language Models

What word fits?

1. “I came here by _____”
2. “The sentence ‘It’s my birthday party’ expresses _____”



A common interpretation of “AI” in the moment: Language Models

What word fits?

1. “I came here by _____”
2. “The sentence ‘It’s my birthday party’ expresses _____”
3. “You need to go through the _____”



A common interpretation of “AI” in the moment: Language Models

What word fits?

1. “I came here by _____”
2. “The sentence ‘It’s my birthday party’ expresses _____”
3. “You need to go through the _____ to reach
‘Untere Sandstrasse’ from the Island in Bamberg.”



A common interpretation of “AI” in the moment: Language Models

What word fits?

1. “I came here by _____”
2. “The sentence ‘It’s my birthday party’ expresses _____”
3. “You need to go through the _____ to reach
‘Untere Sandstrasse’ from the Island in Bamberg.”
 - Language models represent probabilities of words in context.



A common interpretation of “AI” in the moment: Language Models

What word fits?

1. “I came here by _____”
2. “The sentence ‘It’s my birthday party’ expresses _____”
3. “You need to go through the _____ to reach
‘Untere Sandstrasse’ from the Island in Bamberg.”
 - Language models represent probabilities of words in context.
 - Language models can be used for different types of tasks



A common interpretation of “AI” in the moment: Language Models

What word fits?

sentence completion

1. “I came here by _____”
2. “The sentence ‘It’s my birthday party’ expresses _____”
3. “You need to go through the _____ to reach ‘Untere Sandstrasse’ from the Island in Bamberg.”
 - Language models represent probabilities of words in context.
 - Language models can be used for different types of tasks



A common interpretation of “AI” in the moment: Language Models

What word fits?

1. “I came here by _____”
2. “The sentence ‘It’s my birthday party’ expresses _____”
3. “You need to go through the _____ to reach
‘Untere Sandstrasse’ from the Island in Bamberg.”
 - Language models represent probabilities of words in context.
 - Language models can be used for different types of tasks

sentence completion

text classification



A common interpretation of “AI” in the moment: Language Models

What word fits?

1. “I came here by _____”
2. “The sentence ‘It’s my birthday party’ expresses _____”
3. “You need to go through the _____ to reach ‘Untere Sandstrasse’ from the Island in Bamberg.”

sentence completion

text classification

knowledge retrieval

- Language models represent probabilities of words in context.
- Language models can be used for different types of tasks

Inference in Modern Auto-Regressive Language Models





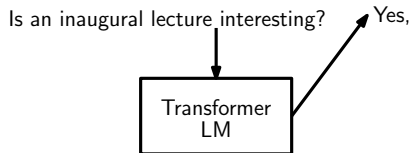
Inference in Modern Auto-Regressive Language Models

Is an inaugural lecture interesting?



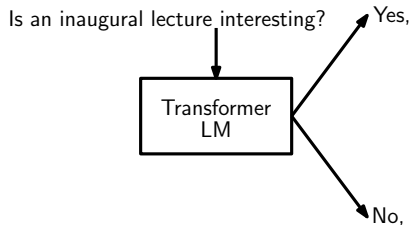


Inference in Modern Auto-Regressive Language Models



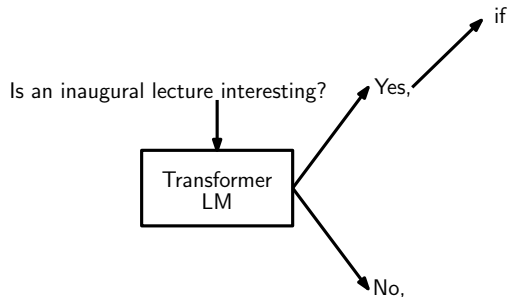


Inference in Modern Auto-Regressive Language Models



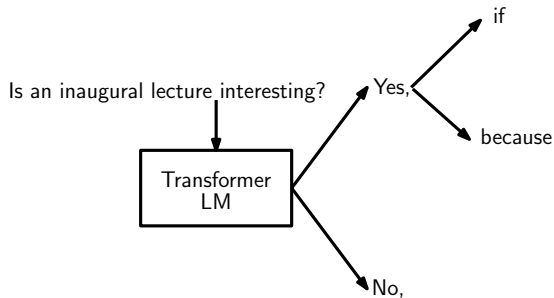


Inference in Modern Auto-Regressive Language Models



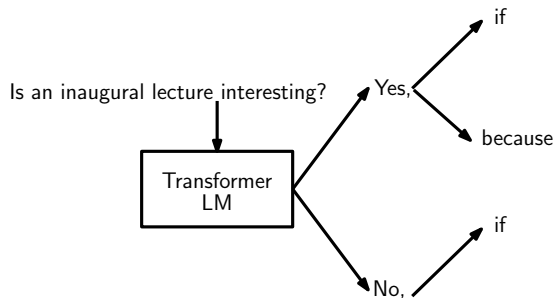


Inference in Modern Auto-Regressive Language Models



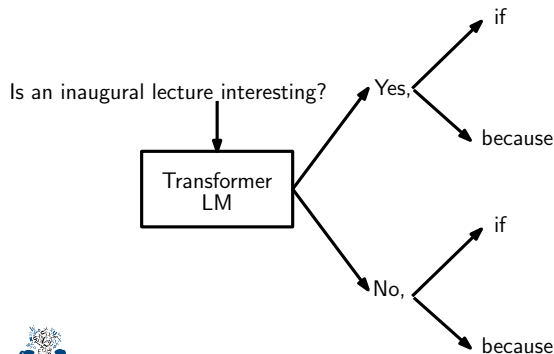


Inference in Modern Auto-Regressive Language Models



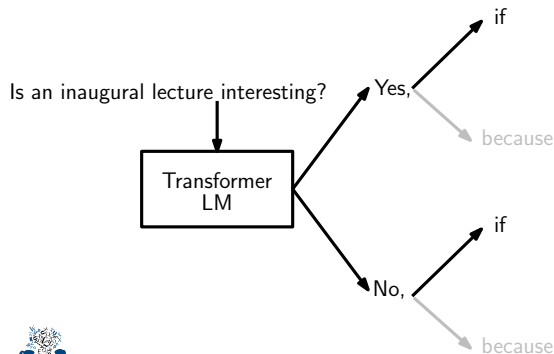


Inference in Modern Auto-Regressive Language Models



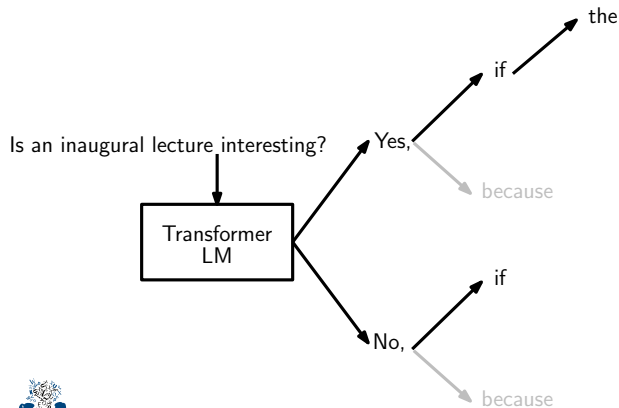


Inference in Modern Auto-Regressive Language Models



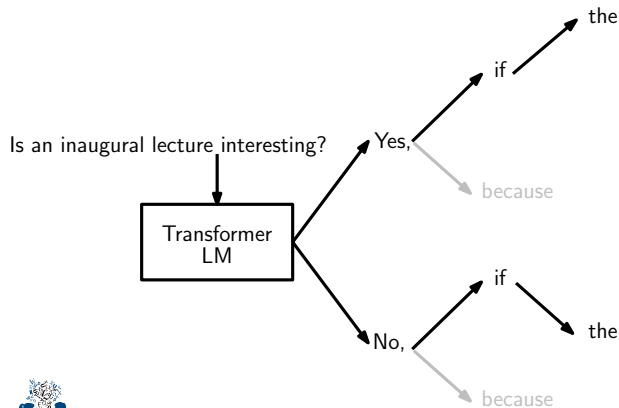


Inference in Modern Auto-Regressive Language Models



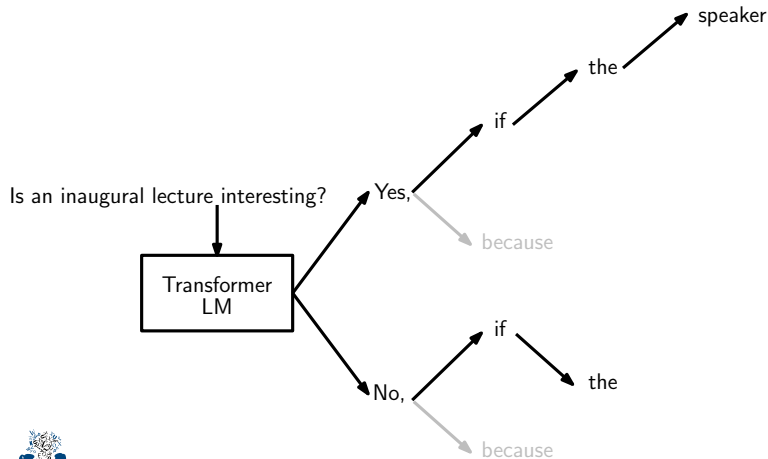


Inference in Modern Auto-Regressive Language Models



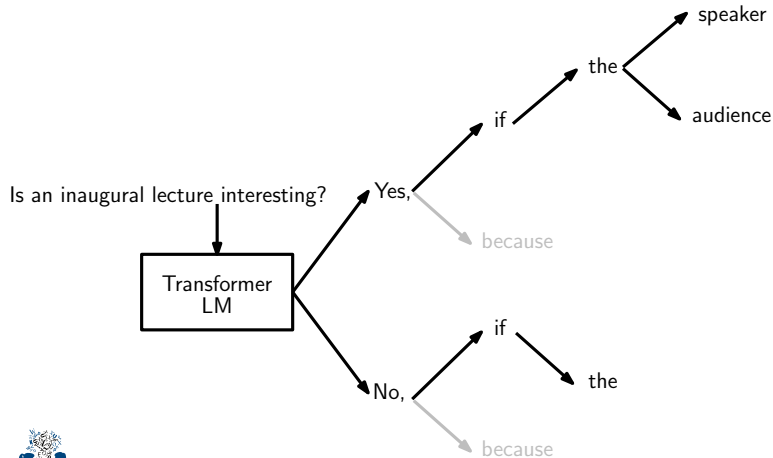


Inference in Modern Auto-Regressive Language Models



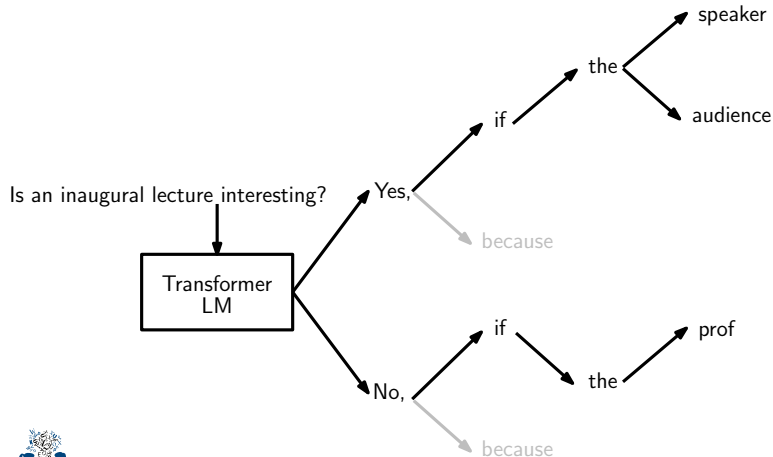


Inference in Modern Auto-Regressive Language Models



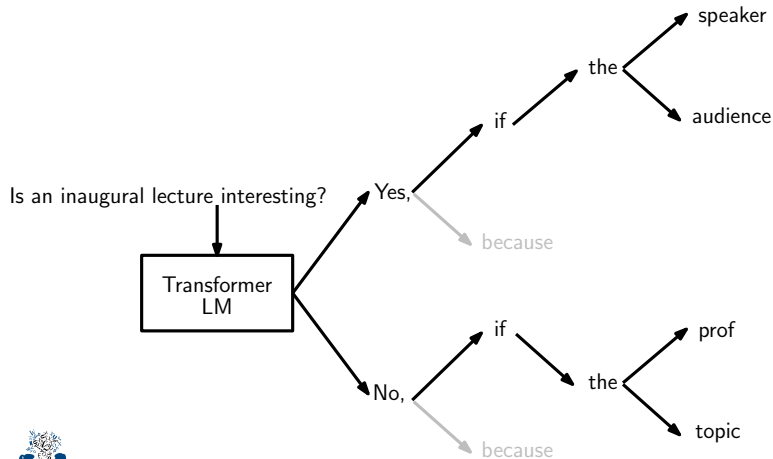


Inference in Modern Auto-Regressive Language Models



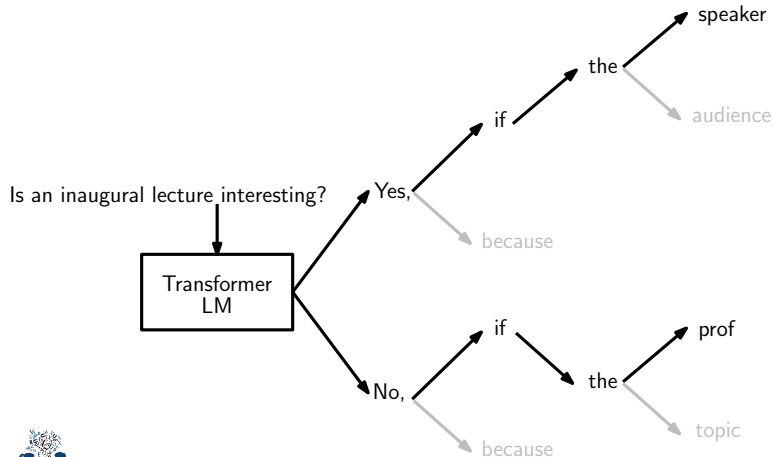


Inference in Modern Auto-Regressive Language Models



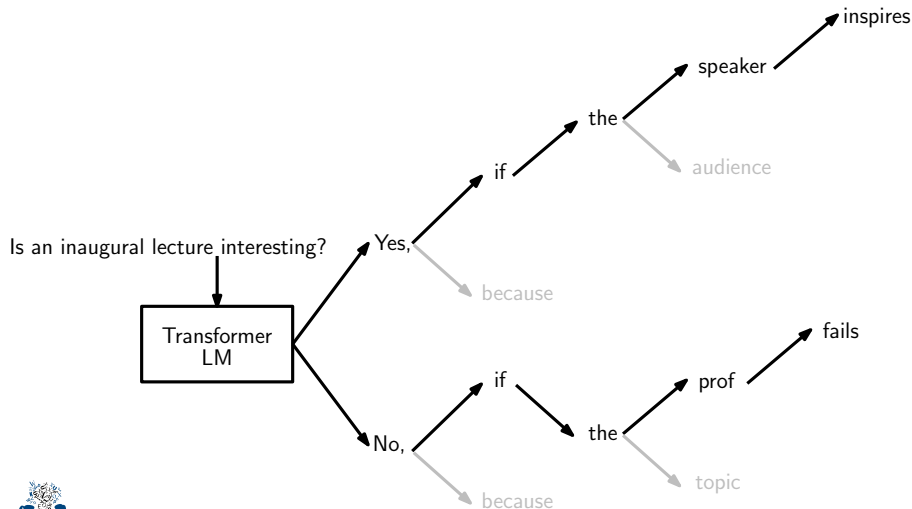


Inference in Modern Auto-Regressive Language Models





Inference in Modern Auto-Regressive Language Models



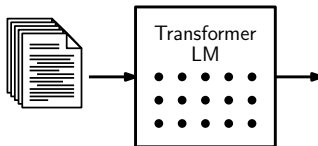
Learning the probabilities in language models





Learning the probabilities in language models

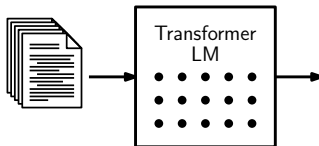
- Large Language Models (LLMs) have **billions of parameters** ...





Learning the probabilities in language models

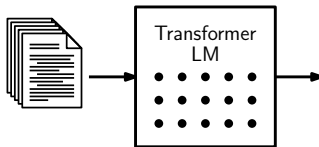
- Large Language Models (LLMs) have **billions of parameters** ...
- ...and see **huge amounts of text** to find values for them.





Learning the probabilities in language models

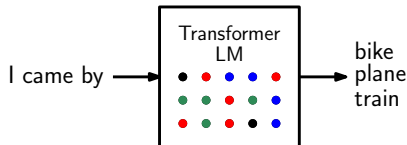
- Large Language Models (LLMs) have **billions of parameters** ...
- ...and see **huge amounts of text** to find values for them.
- The main optimization steps are:





Learning the probabilities in language models

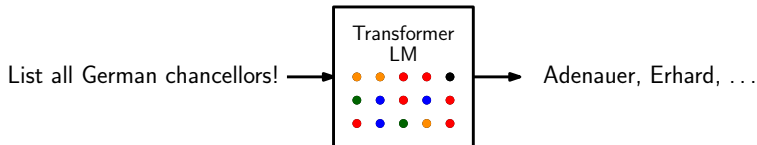
- Large Language Models (LLMs) have **billions of parameters** ...
- ...and see **huge amounts of text** to find values for them.
- The main optimization steps are:
 - **Language modeling objective**





Learning the probabilities in language models

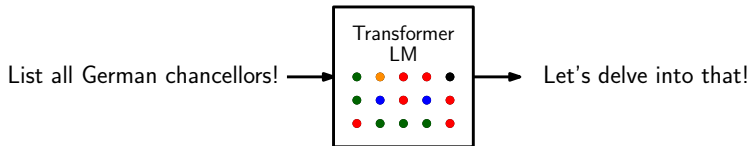
- Large Language Models (LLMs) have **billions of parameters** ...
- ...and see **huge amounts of text** to find values for them.
- The main optimization steps are:
 - **Language modeling objective**
 - **Instruction-tuning**





Learning the probabilities in language models

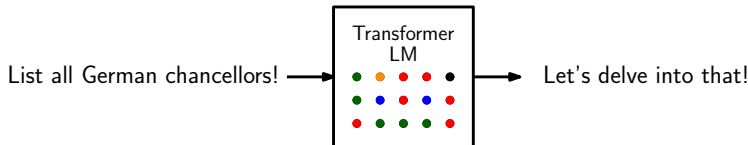
- Large Language Models (LLMs) have **billions of parameters** ...
- ...and see **huge amounts of text** to find values for them.
- The main optimization steps are:
 - **Language modeling objective**
 - **Instruction-tuning**
 - **Alignment through human feedback**





Learning the probabilities in language models

- Large Language Models (LLMs) have **billions of parameters** ...
- ...and see **huge amounts of text** to find values for them.
- The main optimization steps are:
 - **Language modeling objective**
 - **Instruction-tuning**
 - **Alignment through human feedback**
- Finding the best output requires **logical reasoning**, **semantic understanding**, **efficient search**, and **world knowledge**!



What can they do? What can they not do?



What can they do? What can they not do?



- To find out, we need to have data!

What can they do? What can they not do?



- To find out, we need to have data!
- We want to challenge these models.



What can they do? What can they not do?

- To find out, we need to have data!
- We want to challenge these models.
- We are particularly interested in tasks that we consider also challenging for humans.



What can they do? What can they not do?

- To find out, we need to have data!
- We want to challenge these models.
- We are particularly interested in tasks that we consider also challenging for humans.
- We will now look at three cases in which prompting language models does not lead to the desired outcome.

Outline

1 Language Understanding

2 What is “AI”?

3 What we work on

4 Wrap Up

Emotion Detection



F. M. Plaza-del-Arco, M.-T. Martín-Valdivia, and R. Klinger (2022). “Natural Language Inference Prompts for Zero-shot Emotion Classification in Text across Corpora”. In: COLING Fundamentals of Natural Language Processing

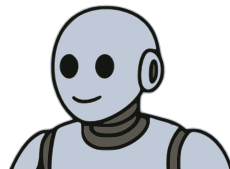
Roman Klinger



Emotion Detection



“You are an expert in emotion annotations.
Which emotion does the author of the following
text feel? Select one from sadness, joy,
anger, fear, surprise, disgust.”

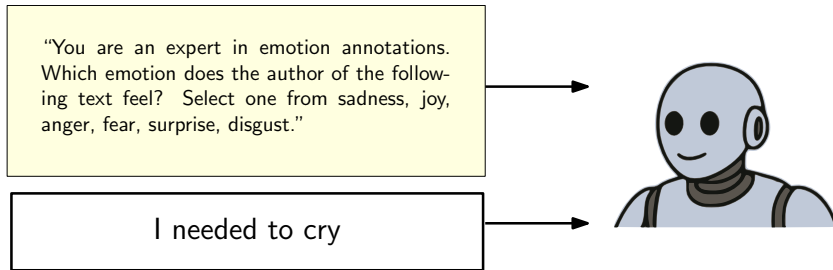


F. M. Plaza-del-Arco, M.-T. Martín-Valdivia, and R. Klinger (2022). “Natural Language Inference Prompts for Zero-shot Emotion Classification in Text across Corpora”. In: COLING
Fundamentals of Natural Language Processing

Roman Klinger



Emotion Detection

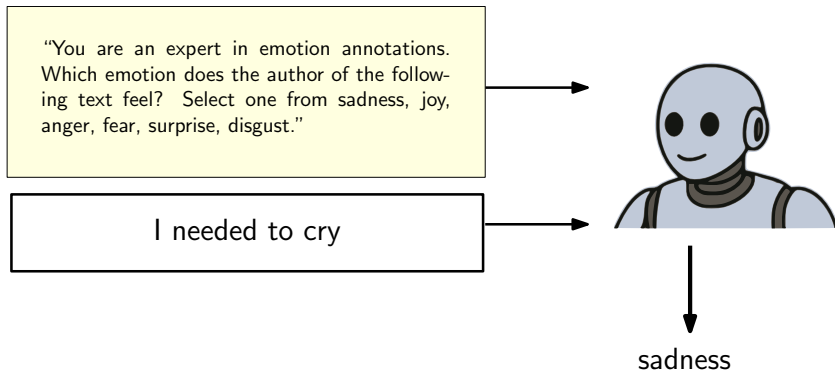


F. M. Plaza-del-Arco, M.-T. Martín-Valdivia, and R. Klinger (2022). "Natural Language Inference Prompts for Zero-shot Emotion Classification in Text across Corpora". In: COLING
Fundamentals of Natural Language Processing

Roman Klinger



Emotion Detection



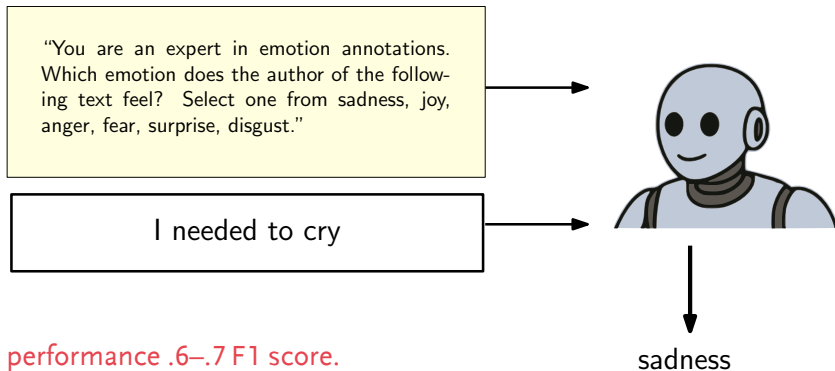
F. M. Plaza-del-Arco, M.-T. Martín-Valdivia, and R. Klinger (2022). "Natural Language Inference Prompts for Zero-shot Emotion Classification in Text across Corpora". In: COLING
Fundamentals of Natural Language Processing

Roman Klinger





Emotion Detection



- Emotion prediction performance .6–.7 F1 score.

F. M. Plaza-del-Arco, M.-T. Martín-Valdivia, and R. Klinger (2022). “Natural Language Inference Prompts for Zero-shot Emotion Classification in Text across Corpora”. In: COLING Fundamentals of Natural Language Processing

Roman Klinger



Emotion Detection in Arguments



L. Greschner and R. Klinger (2025). “Fearful Falcons and Angry Llamas: Emotion Category Annotations of Arguments by Humans and LLMs”. In: NLP4DH

Fundamentals of Natural Language Processing

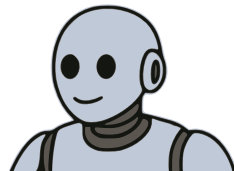
Roman Klinger





Emotion Detection in Arguments

"You are an expert in emotion annotation in arguments. Which emotion does the reader of the following argument likely feel?"



L. Greschner and R. Klinger (2025). "Fearful Falcons and Angry Llamas: Emotion Category Annotations of Arguments by Humans and LLMs". In: NLP4DH

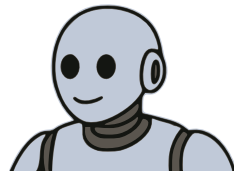




Emotion Detection in Arguments

"You are an expert in emotion annotation in arguments. Which emotion does the reader of the following argument likely feel?"

We should ban plastic bottles, because they harm the environment.

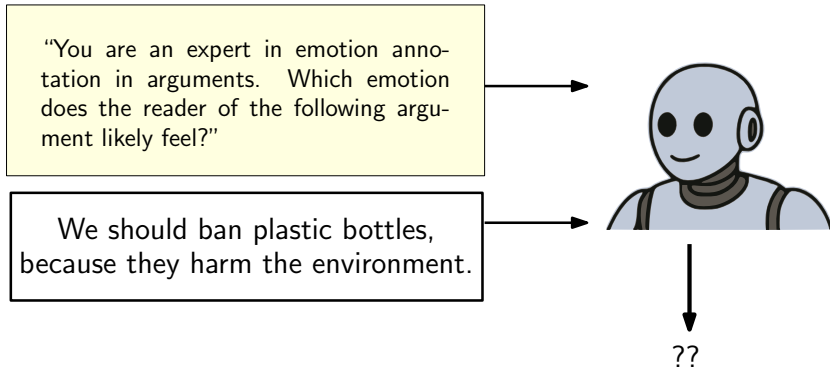


L. Greschner and R. Klinger (2025). "Fearful Falcons and Angry Llamas: Emotion Category Annotations of Arguments by Humans and LLMs". In: NLP4DH





Emotion Detection in Arguments

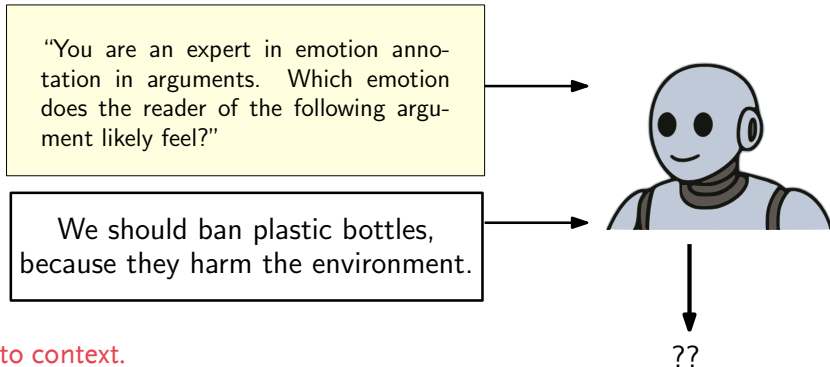


L. Greschner and R. Klinger (2025). "Fearful Falcons and Angry Llamas: Emotion Category Annotations of Arguments by Humans and LLMs". In: NLP4DH





Emotion Detection in Arguments



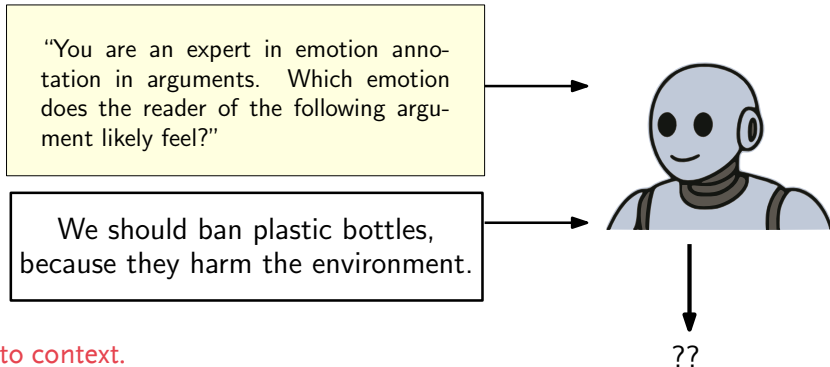
- Models lack access to context.

L. Greschner and R. Klinger (2025). "Fearful Falcons and Angry Llamas: Emotion Category Annotations of Arguments by Humans and LLMs". In: NLP4DH





Emotion Detection in Arguments



- Models lack access to context.
- They tend to predict fear or anger.

L. Greschner and R. Klinger (2025). "Fearful Falcons and Angry Llamas: Emotion Category Annotations of Arguments by Humans and LLMs". In: NLP4DH





How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?

L. Greschner, S. Weber, and R. Klinger (2025). Trust Me, I Can Convince You: The Contextualized Argument Appraisal Framework. under review for LREC 2026. arXiv: 2509.17844 [cs.CL]. URL: <https://arxiv.org/abs/2509.17844>





How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?

- We need to know for whom we make predictions!

L. Greschner, S. Weber, and R. Klinger (2025). Trust Me, I Can Convince You: The Contextualized Argument Appraisal Framework. under review for LREC 2026. arXiv: 2509.17844 [cs.CL]. URL: <https://arxiv.org/abs/2509.17844>





How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?

- We need to know **for whom we make predictions!**
- That is a challenge, we need **annotated arguments** with **information about the annotator!**

L. Greschner, S. Weber, and R. Klinger (2025). Trust Me, I Can Convince You: The Contextualized Argument Appraisal Framework. under review for LREC 2026. arXiv: 2509.17844 [cs.CL]. URL: <https://arxiv.org/abs/2509.17844>





How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?

- We need to know **for whom we make predictions!**
- That is a challenge, we need **annotated arguments** with **information about the annotator!**
- How to get such data?

L. Greschner, S. Weber, and R. Klinger (2025). Trust Me, I Can Convince You: The Contextualized Argument Appraisal Framework. under review for LREC 2026. arXiv: 2509.17844 [cs.CL]. URL: <https://arxiv.org/abs/2509.17844>





How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?

- We need to know **for whom we make predictions!**
- That is a challenge, we need **annotated arguments** with **information about the annotator!**
- How to get such data?
- **We asked people to role play a debate and annotate arguments they read.**

L. Greschner, S. Weber, and R. Klinger (2025). Trust Me, I Can Convince You: The Contextualized Argument Appraisal Framework. under review for LREC 2026. arXiv: 2509.17844 [cs.CL]. URL: <https://arxiv.org/abs/2509.17844>





How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?

- We need to know **for whom we make predictions!**
- That is a challenge, we need **annotated arguments** with **information about the annotator!**
- How to get such data?
- **We asked people to role play a debate and annotate arguments they read.**



L. Greschner, S. Weber, and R. Klinger (2025). Trust Me, I Can Convince You: The Contextualized Argument Appraisal Framework. under review for LREC 2026. arXiv: 2509.17844 [cs.CL]. URL: <https://arxiv.org/abs/2509.17844>



How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?



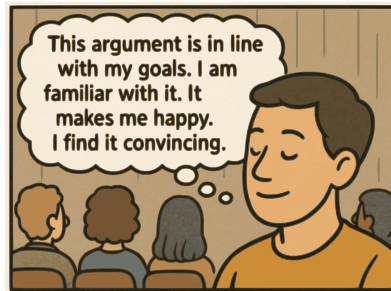


How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?



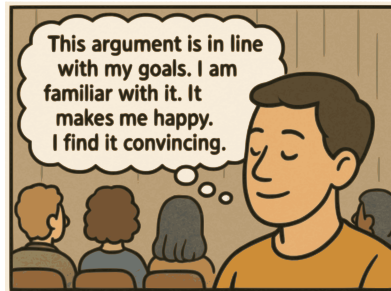


How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?





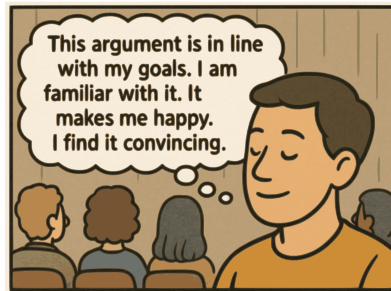
How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?



- We need to develop and evaluate methods to collect data in context.



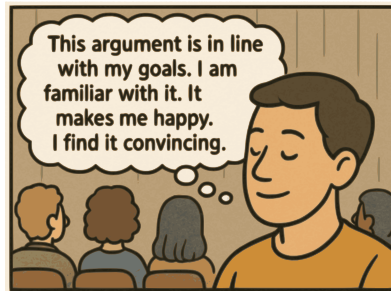
How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?



- We need to develop and evaluate methods to collect data in context.
- Who's the speaker? Who's the listener?



How can we conduct contextualized emotion detection (and convincingness assessment) in arguments?



- We need to develop and evaluate methods to collect data in context.
- Who's the speaker? Who's the listener?
- We need to develop methods to integrate contextual information in computational models.



For whom do models make predictions?



J. Schäfer et al. (2025). "Which Demographics do LLMs Default to During Annotation?" In: ACL



For whom do models make predictions?



- If we don't tell the model for whom it should make a prediction, with whose annotations is it best aligned?

J. Schäfer et al. (2025). "Which Demographics do LLMs Default to During Annotation?" In: ACL



For whom do models make predictions?

- If we don't tell the model for whom it should make a prediction, with whose annotations is it best aligned?



J. Schäfer et al. (2025). "Which Demographics do LLMs Default to During Annotation?" In: ACL



For whom do models make predictions?

- If we don't tell the model for whom it should make a prediction, with whose annotations is it best aligned?



J. Schäfer et al. (2025). "Which Demographics do LLMs Default to During Annotation?" In: ACL

For whom do models make predictions?

- If we don't tell the model for whom it should make a prediction, with whose annotations is it best aligned?



J. Schäfer et al. (2025). "Which Demographics do LLMs Default to During Annotation?" In: ACL





For whom do models make predictions?

- If we don't tell the model for whom it should make a prediction, with whose annotations is it best aligned?
- Models best reconstruct a person's annotation when they are **white, comparably young, and male**.



J. Schäfer et al. (2025). "Which Demographics do LLMs Default to During Annotation?" In: ACL





For whom do models make predictions?

- If we don't tell the model for whom it should make a prediction, with whose annotations is it best aligned?
- Models best reconstruct a person's annotation when they are **white, comparably young, and male**.
- **We need to understand biases and make models work well for everybody.**



J. Schäfer et al. (2025). "Which Demographics do LLMs Default to During Annotation?" In: ACL



Human annotation varies – Should LLM’s annotation also vary?



J. Li, S. Papay, and R. Klinger (2025). “Are Humans as Brittle as Large Language Models?” In: IJCNLP–AAACL



Human annotation varies – Should LLM’s annotation also vary?



- You are an expert in emotion annotation. The label set is $\{LS\}$.
The instance to classify is “The dog ran towards me.”

J. Li, S. Papay, and R. Klinger (2025). “Are Humans as Brittle as Large Language Models?” In: IJCNLP–AAACL





Human annotation varies – Should LLM's annotation also vary?

- You are an expert in emotion annotation. The label set is **{LS}**.
The instance to classify is "The dog ran towards me."
- **anger, fear, joy, disgust, sadness, surprise.**

J. Li, S. Papay, and R. Klinger (2025). "Are Humans as Brittle as Large Language Models?" In: IJCNLP–AAACL



Human annotation varies – Should LLM's annotation also vary?



- You are an expert in emotion annotation. The label set is **{LS}**.
The instance to classify is “The dog ran towards me.”
- **anger, fear, joy, disgust, sadness, surprise.**
- **surprise, sadness, disgust, joy, fear, anger.**

J. Li, S. Papay, and R. Klinger (2025). “Are Humans as Brittle as Large Language Models?” In: IJCNLP–AAACL



Human annotation varies – Should LLM's annotation also vary?



- You are an expert in emotion annotation. The label set is **{LS}**.
The instance to classify is “The dog ran towards me.”
- anger, fear, joy, disgust, sadness, surprise.
- surprise, sadness, disgust, joy, fear, anger.
- angr, feer, joy, disgst, sadnes, suprise.

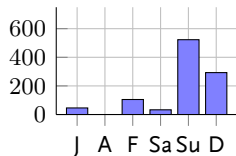
J. Li, S. Papay, and R. Klinger (2025). “Are Humans as Brittle as Large Language Models?” In: IJCNLP–AAACL





Human annotation varies – Should LLM's annotation also vary?

- You are an expert in emotion annotation. The label set is **{LS}**.
The instance to classify is "The dog ran towards me."
- anger, fear, joy, disgust, sadness, surprise.
- surprise, sadness, disgust, joy, fear, anger.
- angr, feer, joy, disgst, sadnes, suprise.



J. Li, S. Papay, and R. Klinger (2025). "Are Humans as Brittle as Large Language Models?" In: IJCNLP–AAACL

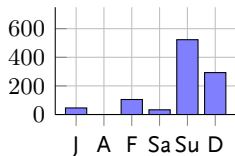




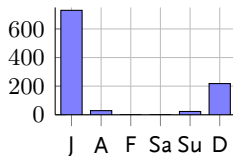
Human annotation varies – Should LLM's annotation also vary?

- You are an expert in emotion annotation. The label set is **{LS}**.
The instance to classify is "The dog ran towards me."

- anger, fear, joy, disgust, sadness, surprise.



- surprise, sadness, disgust, joy, fear, anger.



- angr, feer, joy, disgst, sadnes, suprise.

J. Li, S. Papay, and R. Klinger (2025). "Are Humans as Brittle as Large Language Models?" In: IJCNLP–AAACL

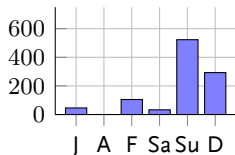




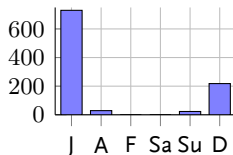
Human annotation varies – Should LLM's annotation also vary?

- You are an expert in emotion annotation. The label set is **{LS}**.
The instance to classify is "The dog ran towards me."

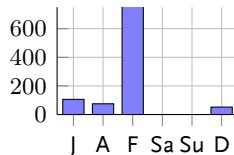
- anger, fear, joy, disgust, sadness, surprise.



- surprise, sadness, disgust, joy, fear, anger.



- angr, feer, joy, disgst, sadnes, suprise.



J. Li, S. Papay, and R. Klinger (2025). "Are Humans as Brittle as Large Language Models?" In: IJCNLP–AAACL

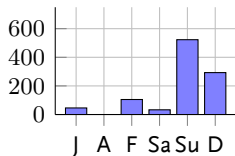




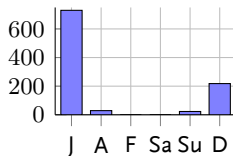
Human annotation varies – Should LLM's annotation also vary?

- You are an expert in emotion annotation. The label set is **{LS}**.
The instance to classify is "The dog ran towards me."

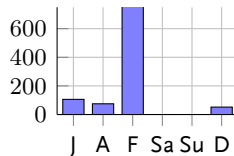
- anger, fear, joy, disgust, sadness, surprise.



- surprise, sadness, disgust, joy, fear, anger.



- angr, feer, joy, disgst, sadnes, suprise.



- Human susceptibility to prompt changes differs from LLM's brittleness.

J. Li, S. Papay, and R. Klinger (2025). "Are Humans as Brittle as Large Language Models?" In: IJCNLP–AAACL

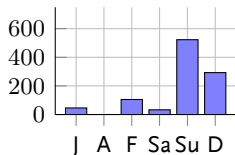




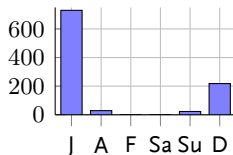
Human annotation varies – Should LLM's annotation also vary?

- You are an expert in emotion annotation. The label set is **{LS}**.
The instance to classify is "The dog ran towards me."

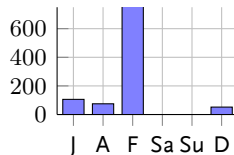
- anger, fear, joy, disgust, sadness, surprise.



- surprise, sadness, disgust, joy, fear, anger.



- anгр, feer, joy, disgst, sadnes, suprise.



- Human susceptibility to prompt changes differs from LLM's brittleness.
- Do we want model's outputs to vary as human's output does?

J. Li, S. Papay, and R. Klinger (2025). "Are Humans as Brittle as Large Language Models?" In: IJCNLP–AAACL

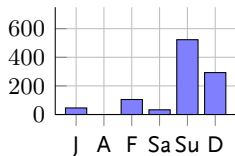




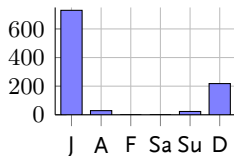
Human annotation varies – Should LLM's annotation also vary?

- You are an expert in emotion annotation. The label set is **{LS}**.
The instance to classify is "The dog ran towards me."

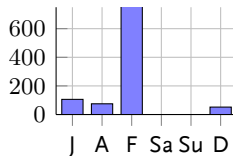
- anger, fear, joy, disgust, sadness, surprise.



- surprise, sadness, disgust, joy, fear, anger.



- anгр, feer, joy, disgst, sadnes, suprise.



- Human susceptibility to prompt changes differs from LLM's brittleness.
- Do we want model's outputs to vary as human's output does?
- If yes, how to achieve that? If no, what should they do?

J. Li, S. Papay, and R. Klinger (2025). "Are Humans as Brittle as Large Language Models?" In: IJCNLP–AACL



Outline

1 Language Understanding

2 What is “AI”?

3 What we work on

4 Wrap Up



What do we do today?

University of Bamberg



Natural Language Processing Research in the Era of AI

Can Computers Read Between the Lines?

Inaugural Lecture, November 18, 2025

Roman Klinger
roman.klinger@uni-bamberg.de

 romanklinger.de  [romanklinger](https://www.linkedin.com/in/romanklinger)
<https://www.bamberg.de/nlproc/>

- What is, and why do we study, Natural Language Processing?
- How does this research field change in the era of "AI"?

Take Home

- What is, and why do we study, Natural Language Processing?
- How does this research field change in the era of “AI”?



Take Home



- What is, and why do we study, Natural Language Processing?
 - NLP is a **research area** that aims at understanding how to **capture language meaning and structure computationally**.
- How does this research field change in the era of “AI”?



Take Home

- What is, and why do we study, Natural Language Processing?
 - NLP is a **research area** that aims at understanding how to **capture language meaning and structure computationally**.
 - **Automatically understanding language** helps to distill **information from large amounts of text**, more than anybody could ever read.
- How does this research field change in the era of "AI"?



Take Home

- What is, and why do we study, Natural Language Processing?
 - NLP is a **research area** that aims at understanding how to **capture language meaning and structure computationally**.
 - **Automatically understanding language** helps to destill **information from large amounts of text**, more than anybody could ever read.
 - Models do not always **behave as expected** and not **for everybody**.
- How does this research field change in the era of "AI"?



Take Home

- What is, and why do we study, Natural Language Processing?
 - NLP is a **research area** that aims at understanding how to **capture language meaning and structure computationally**.
 - **Automatically understanding language** helps to distill **information from large amounts of text**, more than anybody could ever read.
 - Models do not always **behave as expected** and not **for everybody**.
 - To understand what machines should do, we need to **understand human language in context**.
- How does this research field change in the era of “AI”?



Take Home

- What is, and why do we study, Natural Language Processing?
 - NLP is a **research area** that aims at understanding how to **capture language meaning and structure computationally**.
 - **Automatically understanding language** helps to distill **information from large amounts of text**, more than anybody could ever read.
 - Models do not always **behave as expected** and not **for everybody**.
 - To understand what machines should do, we need to **understand human language in context**.
- How does this research field change in the era of “AI”?
 - LLMs do **not** (yet) just solve tasks **“off the shelf”** ...



Take Home

- What is, and why do we study, Natural Language Processing?
 - NLP is a **research area** that aims at understanding how to **capture language meaning and structure computationally**.
 - **Automatically understanding language** helps to destill **information from large amounts of text**, more than anybody could ever read.
 - Models do not always **behave as expected** and not **for everybody**.
 - To understand what machines should do, we need to **understand human language in context**.
- How does this research field change in the era of “AI”?
 - LLMs do **not** (yet) just solve tasks **“off the shelf”** ...
 - ...but they show **promising results** for many tasks.



Take Home

- What is, and why do we study, Natural Language Processing?
 - NLP is a **research area** that aims at understanding how to **capture language meaning and structure computationally**.
 - **Automatically understanding language** helps to distill **information from large amounts of text**, more than anybody could ever read.
 - Models do not always **behave as expected** and not **for everybody**.
 - To understand what machines should do, we need to **understand human language in context**.
- How does this research field change in the era of “AI”?
 - LLMs do **not** (yet) just solve tasks **“off the shelf”** ...
 - ...but they show **promising results** for many tasks.
 - Systems relying on more **traditional methods** often work **better**.





- Are we bread scholars?



- Are we **bread scholars**?
 - We do teach how to practically build systems that can understand language. (so...; Yes?)



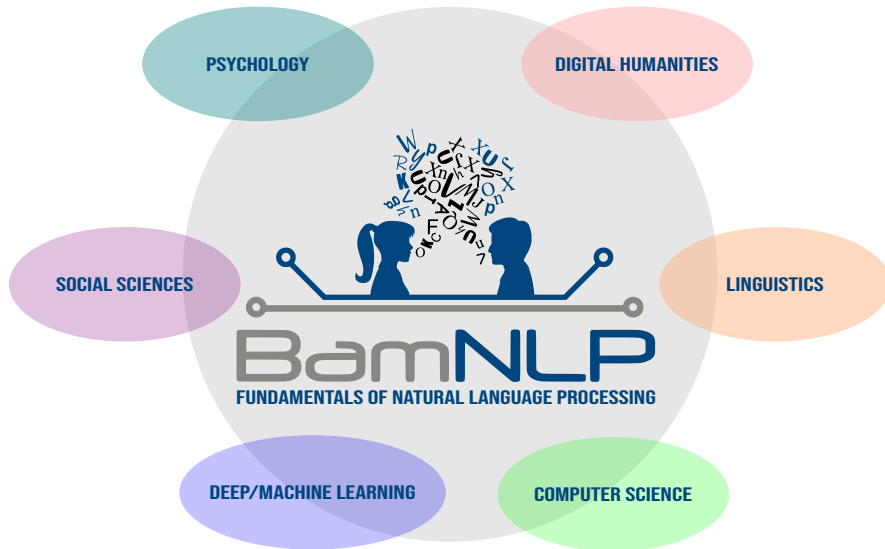
- Are we **bread scholars**?
 - We do teach how to practically build systems that can understand language. (so...; Yes?)
 - To advance the field, however, a **deeper understanding** of human **expectations**, their **behaviour**, their **properties that influence language interpretation and production** is needed. (so...; No?)



- Are we **bread scholars**?
 - We do teach how to practically build systems that can understand language. (so...; Yes?)
 - To advance the field, however, a **deeper understanding** of human **expectations**, their **behaviour**, their **properties that influence language interpretation and production** is needed. (so...; No?)
- Can we be **philosophical minds**?



- Are we **bread scholars**?
 - We do teach how to practically build systems that can understand language. (so...; Yes?)
 - To advance the field, however, a **deeper understanding** of human **expectations**, their **behaviour**, their **properties that influence language interpretation and production** is needed. (so...; No?)
- Can we be **philosophical minds**?
 - **We need to be!**





Thank you for
your attention.
Questions? Remarks?



Funded by
DFG Deutsche
Forschungsgemeinschaft
German Research Foundation

Thanks to:



- All research groups I was part of so far and all collaborators.
- Joy Kearney, Mona Wirth and Romy Schatzschneider for the organization.
- All of you for your interest!
- Please reach out if you want to talk, chat, discuss, meet us, drink coffee, work with us, collaborate, ...



<https://tinyurl.com/bambergtranslate>



Natural Language Processing Research in the Era of AI

Can Computers Read Between the Lines?

Inaugural Lecture, November 18, 2025

Roman Klinger

roman.klinger@uni-bamberg.de

 [romanklinger.de](https://twitter.com/romanklinger)  [romanklinger](https://www.linkedin.com/in/romanklinger)

<https://www.bamberg.de/nlproc/>

<https://www.romanklinger.de/>