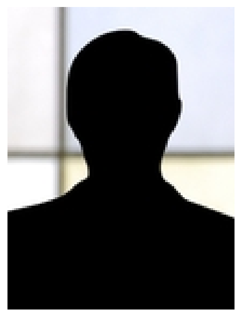
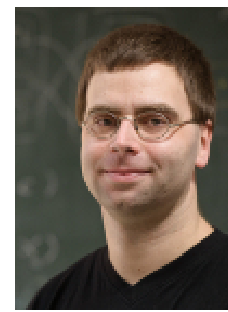
Towards Understanding and Answering Comparative Questions

Published at WSDM'22

March 11, 2022





Alexander Bondarenko[1] Yamen Ajjour[1] Valentin Dittmar[1] Niklas Homann[1] Pavel Braslavski[2,3] Matthias Hagen[1]

[1]Martin-Luther-Universität Halle-Wittenberg, [2]Ural Federal University, [3]HSE University

www.webis.de

Webis Flash Talks; Virtual Event.

Towards Understanding and Answering Comparative Questions

Motivation

Direct:	**Is a cat or a dog a better friend?**	
	object: cat; object: dog; predicate: better; aspect: friend	
Indirect:	**What pet is the best friend?**	
	object: pet; predicate: best; aspect: friend	
Without aspect:	**Who is better, a cat or a dog?**	
	predicate: better; object: cat; object: dog	

Pro obj. 1: *Cats can be quite affectionate and attentive, and thus are good friends.*

Pro obj. 2: *Cats are less faithful than dogs.*

Towards Understanding and Answering Comparative Questions

Motivation

Direct: **`Is a cat or a dog a better friend?`**
cat: object, dog: object, better: predicate, friend: aspect

Indirect: **`What pet is the best friend?`**
pet: object, best: predicate, friend: aspect

Without aspect: **`Who is better, a cat or a dog?`**
better: predicate, cat: object, dog: object

Pro obj. 1: *Cats can be quite affectionate and attentive, and thus are good friends.*
Pro obj. 2: *Cats are less faithful than dogs.*

- Dataset: comparative questions with objects, aspects, and answers' stances.
- Classifiers for comparative and subjective comparative questions.
- Identifying objects, aspects, and predicates.
- Stance detector for answers.

Towards Understanding and Answering Comparative Questions

Comprative Question Parsing

`Is a cat or a dog a better friend?`

Is a	cat	or a	dog	a	better	friend?
	object		object		predicate	aspect

- Cascading ensemble recalls 71% of comparative questions at prec. of 1.0.
- Identifying subjective questions (RoBERTa): F1 0.95.
- Baseline: BiLSTM with 300-dimensional GloVe embeddings [Arora et al.; CIKM'17].

	F1 scores			
Classifier	**Object**	**Aspect**	**Predicate**	**None**
BiLSTM	0.80	0.52	0.85	**0.98**
RoBERTa	**0.93**	**0.80**	**0.98**	0.94

Towards Understanding and Answering Comparative Questions

Answer Stance Detection

`Is a cat or a dog a better friend?`

Pro obj. 1: *Cats can be quite affectionate and attentive, and thus are good friends.*

Towards Understanding and Answering Comparative Questions

Answer Stance Detection

`Is a OBJECT 1 or a OBJECT 2 a better friend?`

Pro obj. 1: *OBJECT 1 can be quite affectionate and attentive, and thus are good friends.*

- Most effective classifier RoBERTa.
- Comparison objects are masked in questions and answers.
- Add a sentiment prompt: OBJECT 1 *is better*.
- Input: OBJECT 1 *is better* `[SEP]` ANSWER.
- Highest accuracy on 4 labels (pro object 1 / 2, neutral, no stance) 0.63.

Towards Understanding and Answering Comparative Questions

Conclusions

- Dataset: comparative questions with objects, aspects, and answers' stances.
- Classifiers for comparative questions, objects, aspects, and predicates.
- Stance detector for potential answers.

Towards Understanding and Answering Comparative Questions

Conclusions

- Dataset: comparative questions with objects, aspects, and answers' stances.
- Classifiers for comparative questions, objects, aspects, and predicates.
- Stance detector for potential answers.

thank you!

Propaganda Technique Detection using Connotation Frames

-Vinaykumar Budanurmath

Thesis Supervisor: Jun.-Prof. Dr. Henning Wachsmuth
Second supervisor: Jun.-Prof. Sebastian Peitz
Guide: PhD student. Mr. Wei Fan Chen

PADERBORN UNIVERSITY

1. Problem Statement

a) On November 25, a soldier veered his jeep into a crowded market and killed three civilians.
b) On November 25, a soldier's jeep veered into a crowded market, causing three civilian deaths.

(a) Blames the soldier by associating the verbs **killed** and **veered** with the soldier

(b) The blame is shifted from the soldier to the jeep.

- One propaganda many emotions!
- Propaganda meaning and its characteristics:
  - Is a form of persuasion
  - Aims to influence the reader's opinions
    - By targeting the emotions
  - Deceptive

1. Problem Statement continued…

- Many propaganda techniques
  - Corpus contained 19 techniques
- Example
  - *"He warned that the danger was not over".*

    Technique: *Appeal to fear-prejudice*
  - *"....a striking blow against the freedom."*

    Technique: Loaded language
- GOAL – Find propaganda techniques in a propaganda statement

Appeal_to_Authority
Appeal_to_fear-prejudice
Bandwagon
Black-and-White_Fallacy
Causal_Oversimplification
Doubt
Exaggeration
Flag-Waving
Labeling
Loaded_Language
Minimisation
Name_Calling
Red_Herring
Reductio_ad_hitlerum
Repetition
Slogans
Straw_Men
Thought-terminating_Cliches
Whataboutism

2. Motivation

- Connotations in text
  - *a feeling or idea that is suggested by a particular word* (Source: Cambridge dictionary)

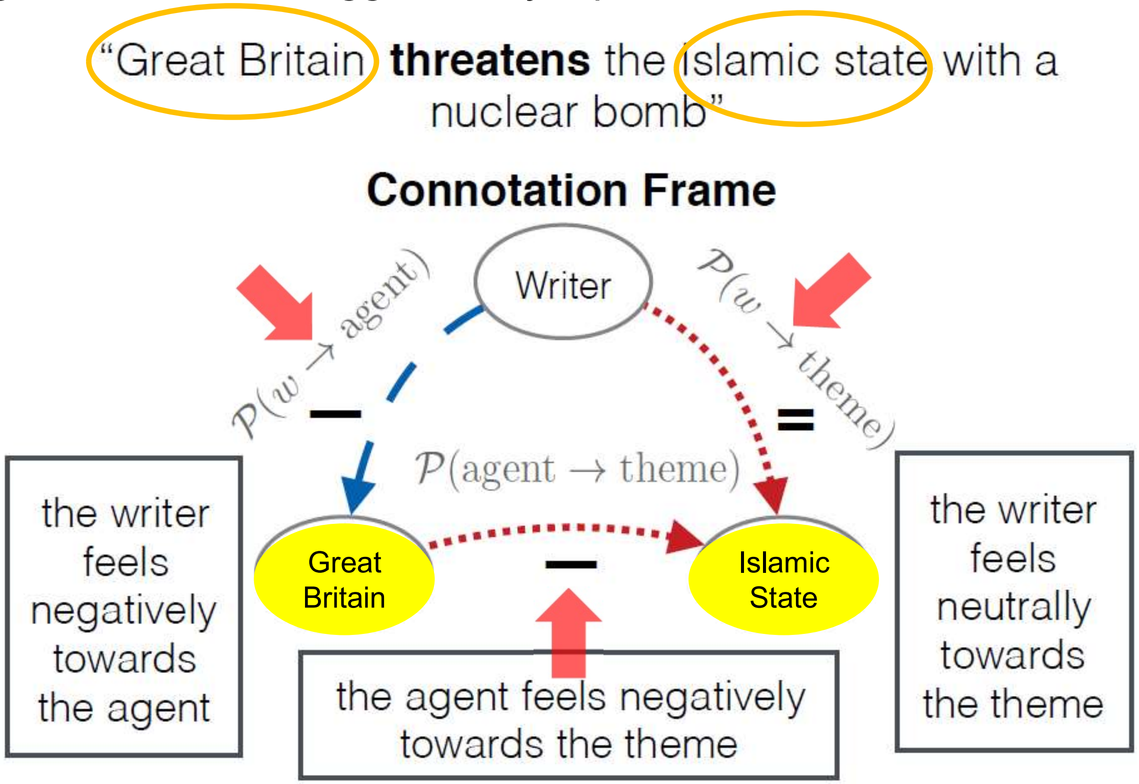


Fig: Connotation frame: Source: Volkova and Jang (2018)

- The connotation frames were introduced by researchers Rashkin et al.,

2. Motivation continued….

verb	P(wo)	P(ws)	P(so)	E(o)	E(s)	V(o)	V(s)	S(o)	S(s)	P(ro)	P(rs)	P(os)
have	0.37	0.3	0.47	0.07	0.2	0.47	0.6	0.07	0.37	0.4	0.37	0.07
say	0.0	0.17	0.07	0.13	0.07	0.07	1.0	0.0	0.1	0.03	0.27	0.0
take	0.5	0.6	0.8	0.13	0.6	0.47	0.6	0.03	0.67	0.5	0.57	0.07
go	0.07	0.13	0.2	0.0	0.07	-0.07	0.47	-0.1	0.03	0.07	0.07	-0.1

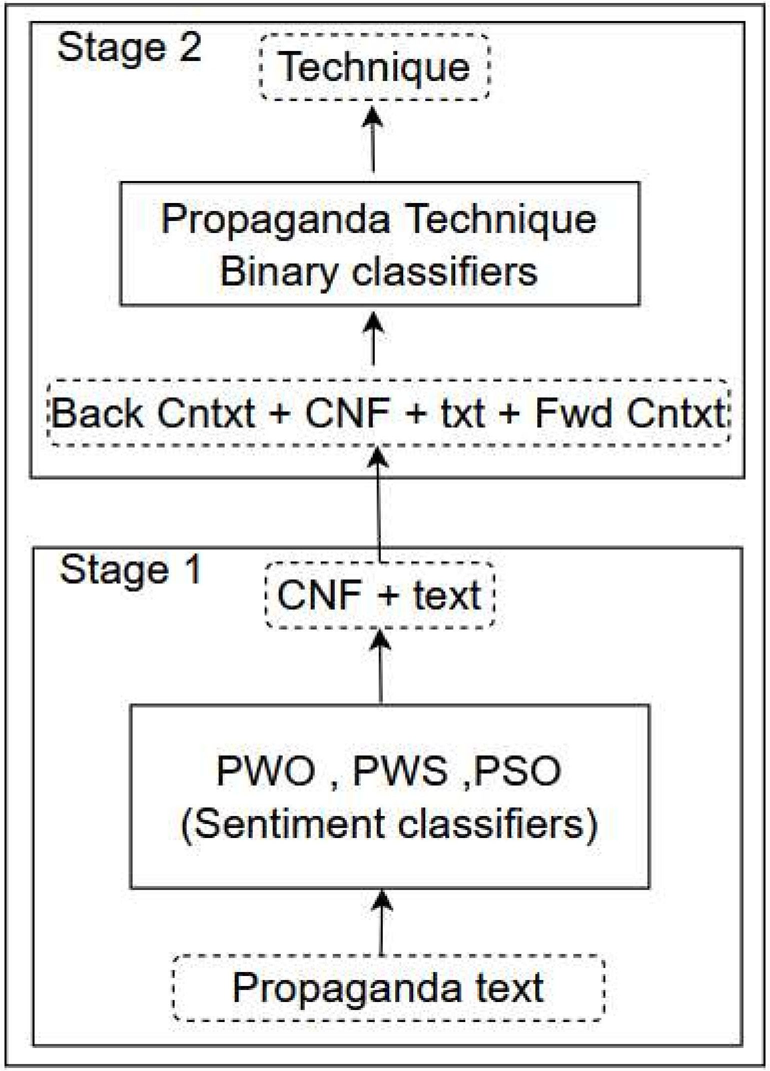
- [-1.0, -0.25) : -/Negative
- [-.25, 0.25] : =/Neutral
- (0.25, 1.0] : +/Positive
  - Example: "threaten"
  - PWO – Sentiment from Writer to Object: 0.03 (Neutral)
  - PWS – Sentiment from Writer to Subject: - 0.3 (Negative)
  - PSO – Sentiment from Subject to Object: - 0.63 (Negative)

[CNF_START] PWO NEU PWS NEG PSO NEG [CNF_END] + <Propaganda text>

- Objectives:
  - Will connotations be helpful in propaganda technique detection?
  - What are the sentiments involved in each technique?

3. Approach

Machine Learning Model used in both stages: Transformers with BERT Language model with 10 epochs in training

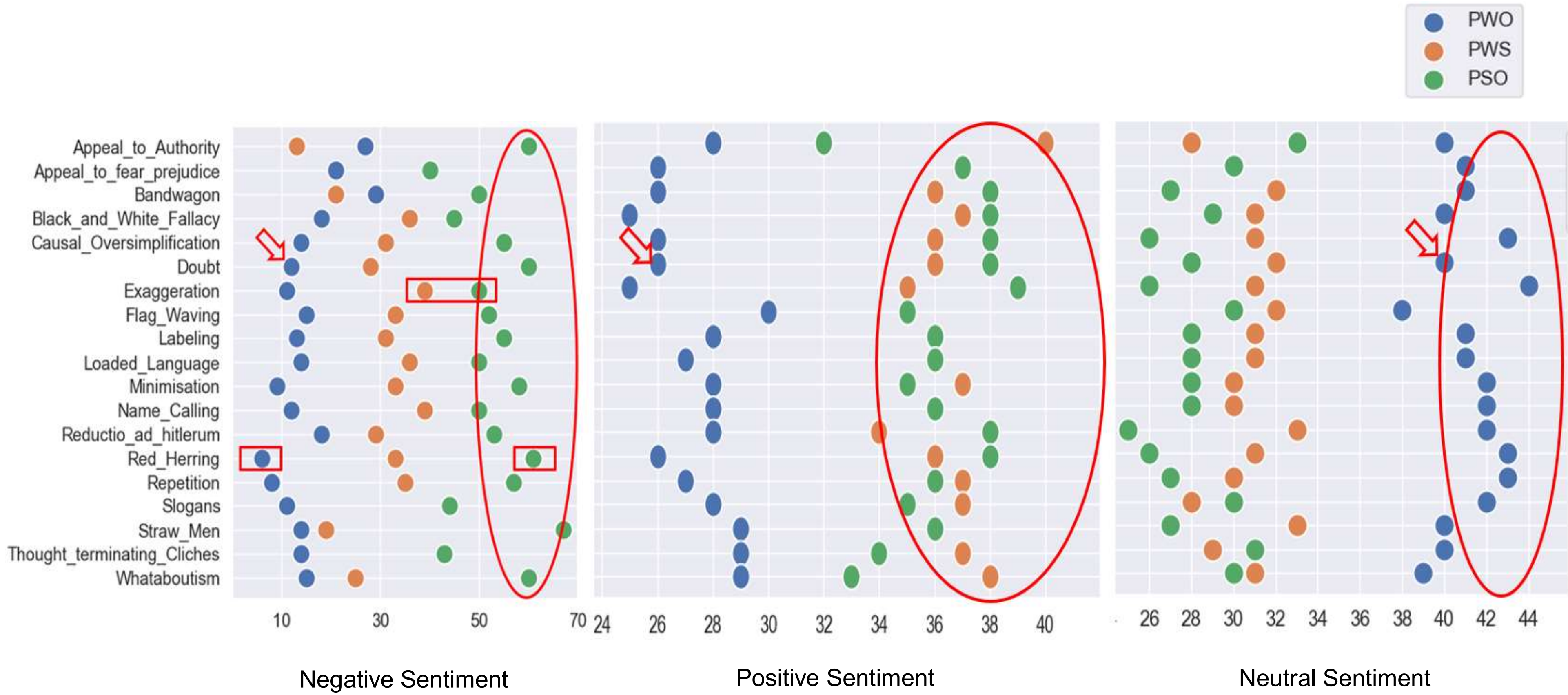


Stage 2
Propaganda Technique Classifier (PTC) Model
Identifies the technique

Stage 1
Computational Connotation Frame (CCF) Model
Helps us understand the sentiments

4. Model Results

- CCF model results – sentiment classifier results



Axes Legend
X- Axis: Percentages of number of statements
Y- Axis: Propaganda techniques

4. Model Results

- Propaganda Technique Classifier (PTC) model results
- F1 score – 0.45
- Top score – 0.63

5. Future Work

- Future work:
  - Use the expanded set of Rashkin et al result set
  - Use the sentiment perspectives in summarization.

Thank you

Distilling the Finest Memes from Web Archive Data Using CPU and GPU Clusters

Niklas Deckers
Leipzig University

Webis Flash Talks 2022

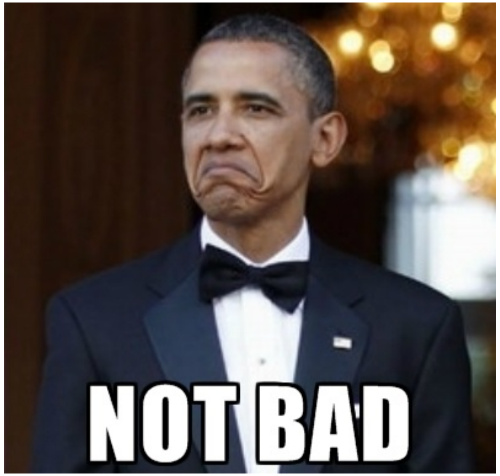
Dummy Task: Identifying Meme Images from the Internet Archive

- Given: Petabytes of WARC files with web archive data from the Internet Archive
- Goal: Extract meme images
- Problem: Don’t know the location of memes within the WARC files

Dummy Task: Identifying Meme Images from the Internet Archive

- Given: Petabytes of WARC files with web archive data from the Internet Archive
- Goal: Extract meme images
- Problem: Don't know the location of memes within the WARC files

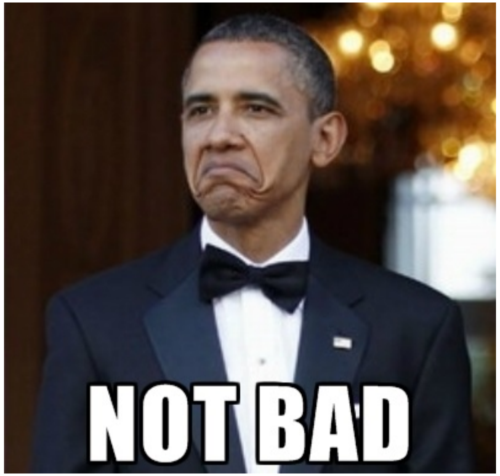
- Approach: Use any pretrained meme image classifier from Github
(claims 93% accuracy on test data)



Dummy Task: Identifying Meme Images from the Internet Archive

- Given: Petabytes of WARC files with web archive data from the Internet Archive
- Goal: Extract meme images
- Problem: Don't know the location of memes within the WARC files

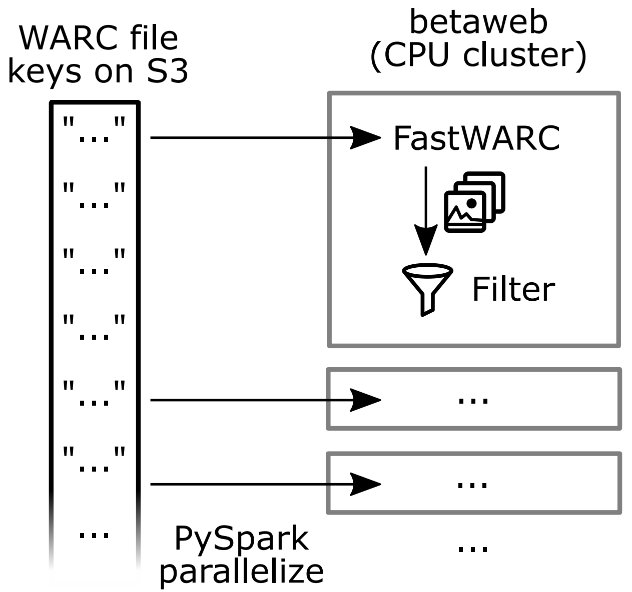
- Approach: Use any pretrained meme image classifier from Github
(claims 93% accuracy on test data)



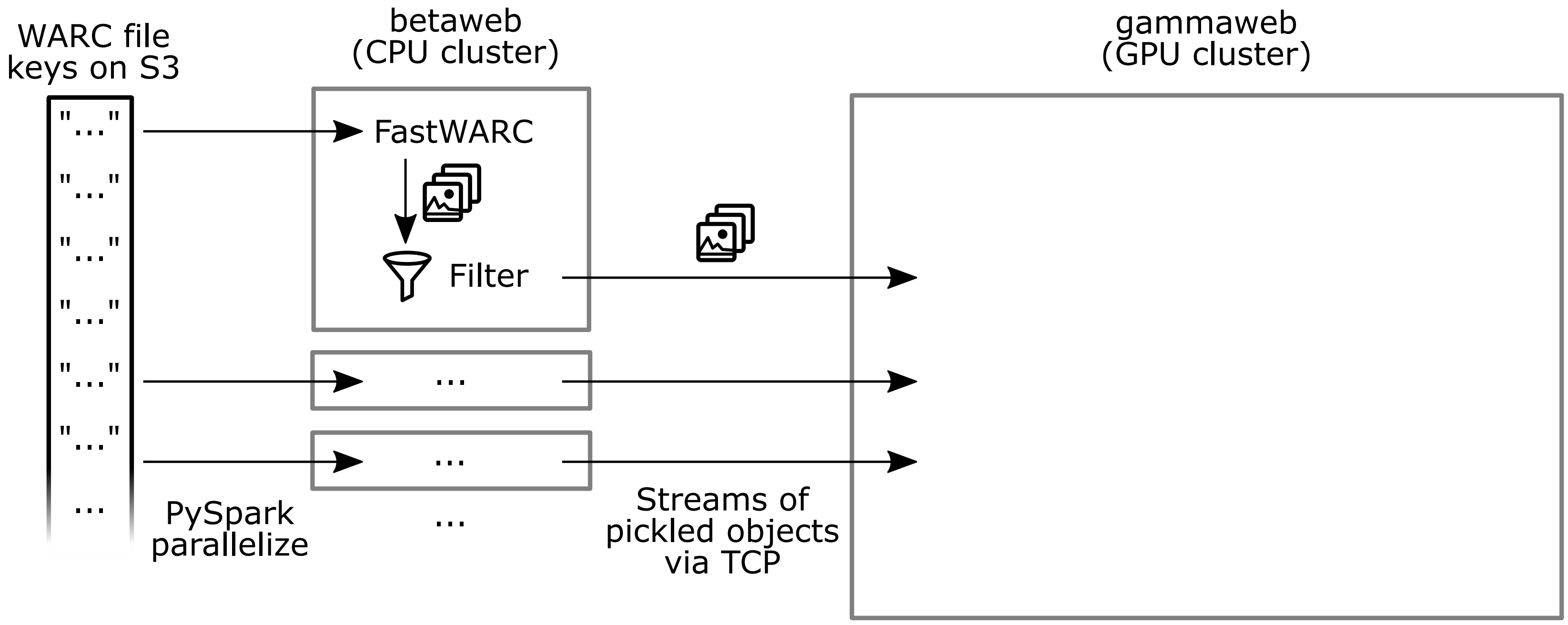
Conflicting Hardware Requirements

- Iterating through WARC files is CPU-intensive
  and should be parallelized on a CPU cluster (betaweb)
- Image classification works best on GPU (not available on betaweb)
- Solution:
  - WARC extraction and preprocessing on CPU cluster
  - Prediction (network inference) on GPU

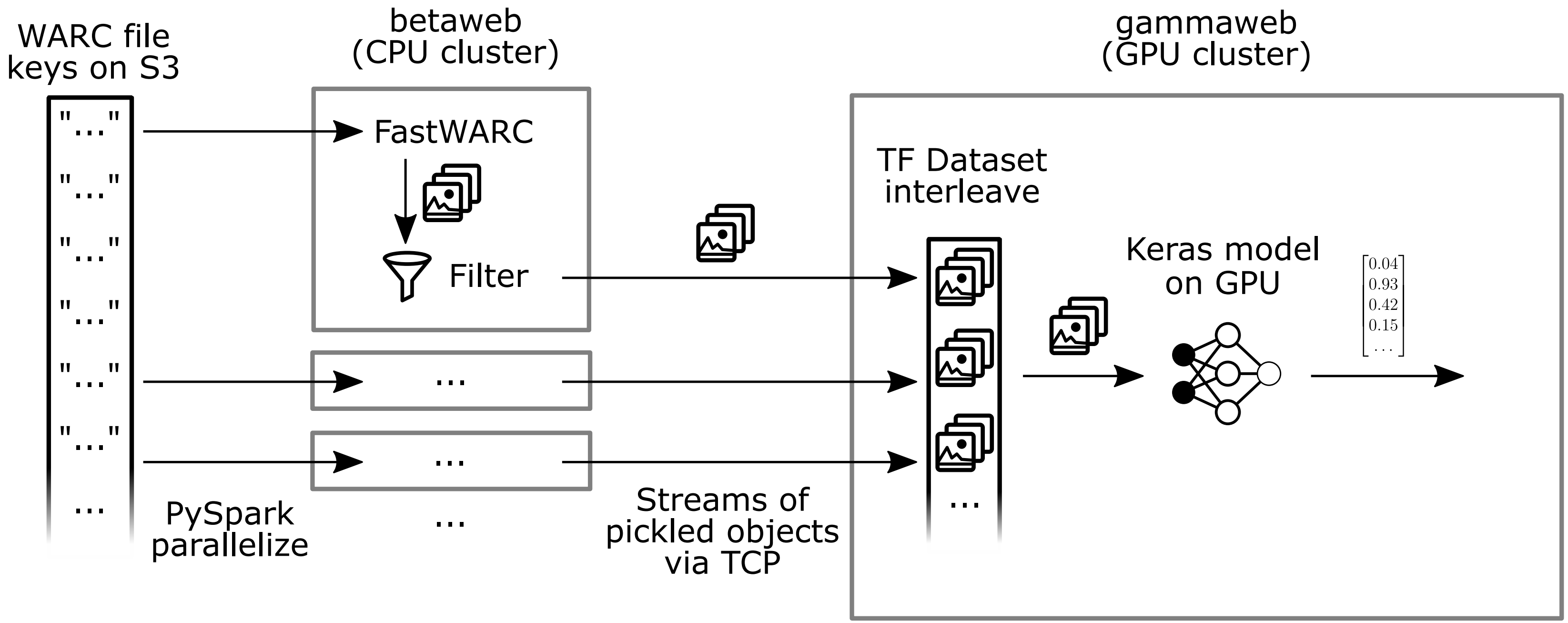
Proposed web-archive-keras Pipeline Architecture



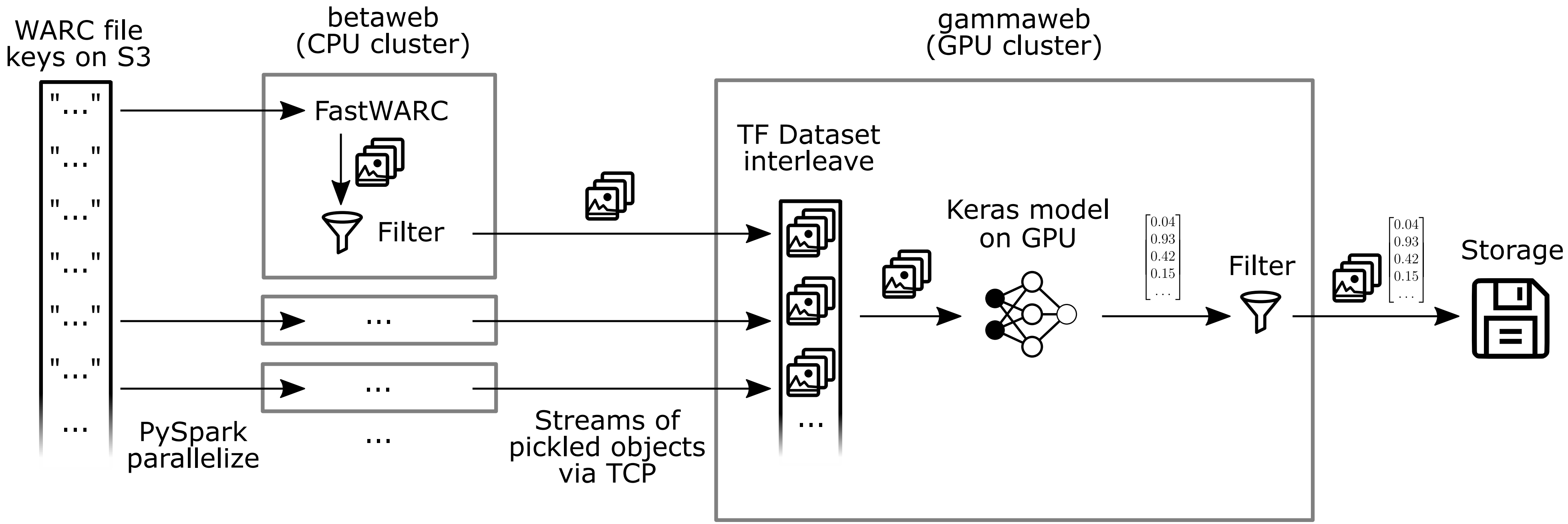
Proposed web-archive-keras Pipeline Architecture



Proposed web-archive-keras Pipeline Architecture



Proposed web-archive-keras Pipeline Architecture



Performance

- Troughput of the pipeline with 100 spark instances and one GPU: 50 images per second
- Quick evaluation on approx. 1000 images:

Performance

- Troughput of the pipeline with 100 spark instances and one GPU:
  50 images per second
- Quick evaluation on approx. 1000 images:

	Without filtering
Funny images	0.14%
Images with high meme potential	0.28%
Established memes	0.07%

Performance

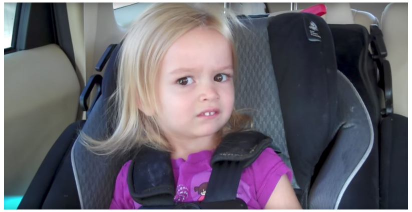
- Troughput of the pipeline with 100 spark instances and one GPU: 50 images per second
- Quick evaluation on approx. 1000 images:

	Without filtering
Funny images	0.14%
Images with high meme potential	0.28%
Established memes	0.07%
Adult/NSFW images	10.21%

Performance

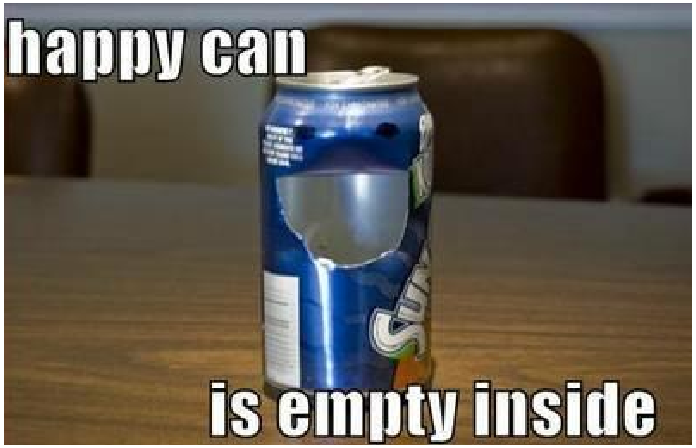
- Troughput of the pipeline with 100 spark instances and one GPU: 50 images per second
- Quick evaluation on approx. 1000 images:

	Without filtering	**With filtering by model prediction (score >0.95)**
Funny images	0.14%	1.94%
Images with high meme potential	0.28%	0.78%
Established memes	0.07%	0.19%
Adult/NSFW images	10.21%	1.94%

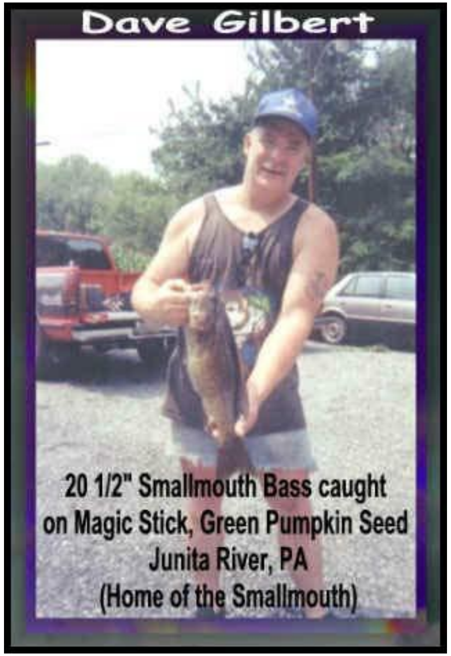


Error Analysis: The Classifier’s Finest Memes

Established Memes



Error Analysis: The Classifier’s Finest Memes
Looks Like Meme but Is Not



Conclusion

- Code for pipeline with examples (image and text classification) is available
- Deployment instructions in Webis knowledgebase
- Easily adaptable to more models
- Planned features:
  - Model training
  - Multi-GPU
  - Combining text and image information

Conclusion

- Code for pipeline with examples (image and text classification) is available
- Deployment instructions in Webis knowledgebase
- Easily adaptable to more models
- Planned features:
  - Model training
  - Multi-GPU
  - Combining text and image information
- Takeaway: Web archive data is often out of distribution for pretrained models

RULES TO THINK OUTSIDE THE BOX

A DATASET OF RIDDLES

Flashtalk by Theresa Elstner

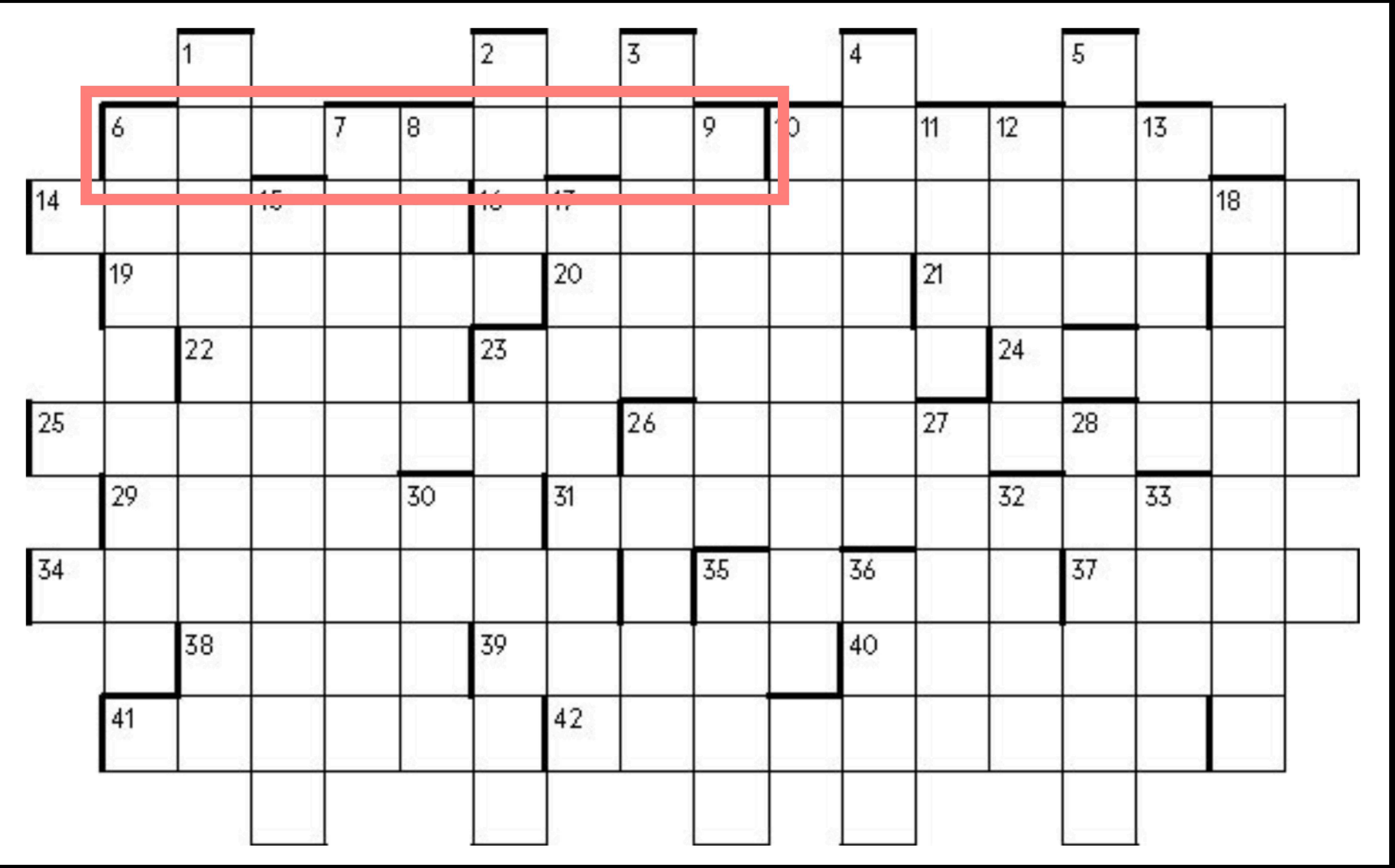
THINK OUTSIDE THE BOX RIDDLE

ZEIT MAGAZIN - UM DIE ECKE GEDACHT

6 HORIZONTAL

Aufs Navi zu hören ist ein …, sich das … zu ersparen

(Translates to: Following a navigation system's route is a …, to prevent …)



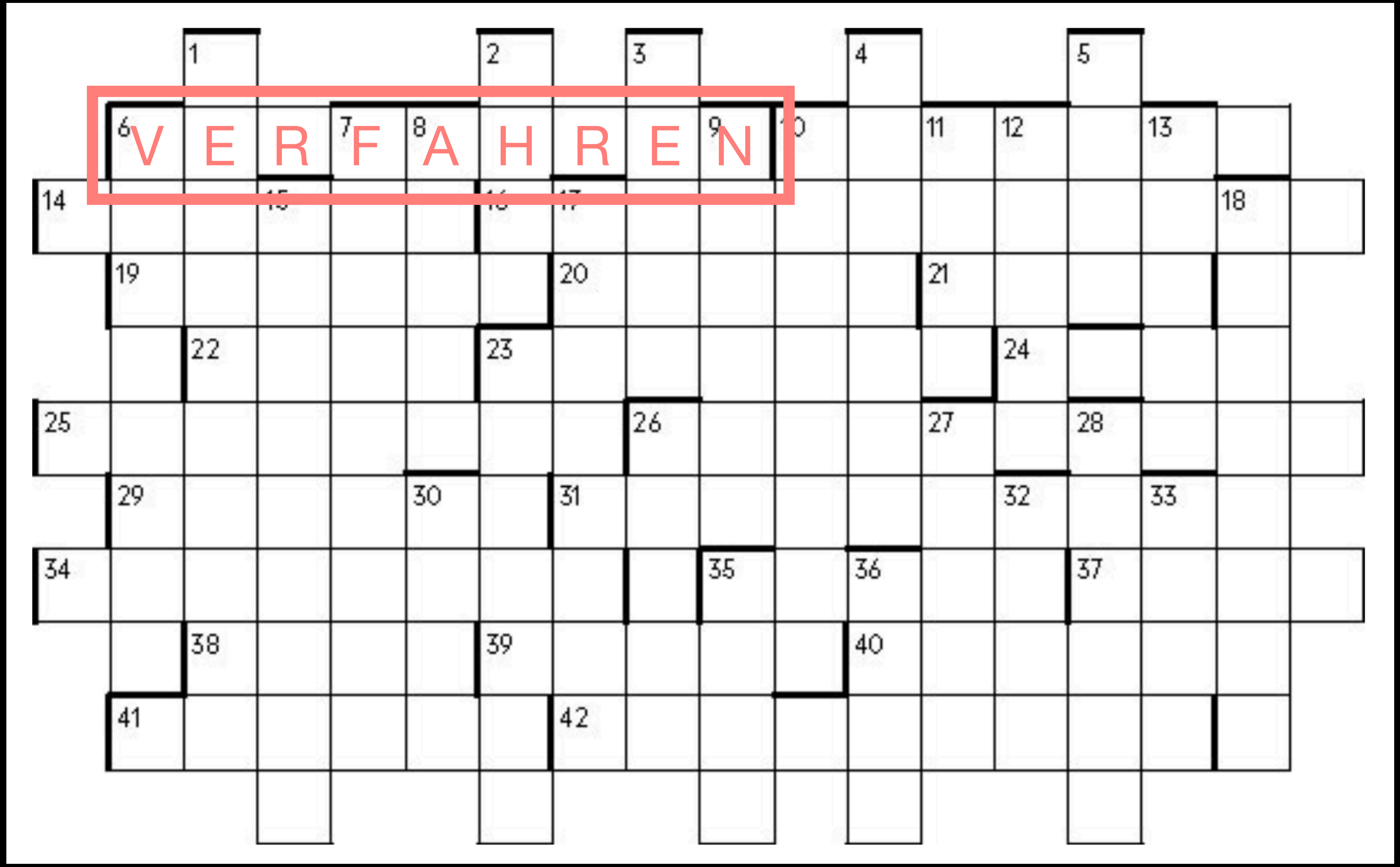
ZEIT MAGAZIN RÄTSEL

What is that?

ANSWER

Aufs Navi zu hören ist ein Verfahren, sich das Verfahren zu ersparen

(Translates to: Following a navigation system's route is a method, to prevent getting lost)



RULE 1

Use multiple meanings of the same word in different contexts and leave out the word.

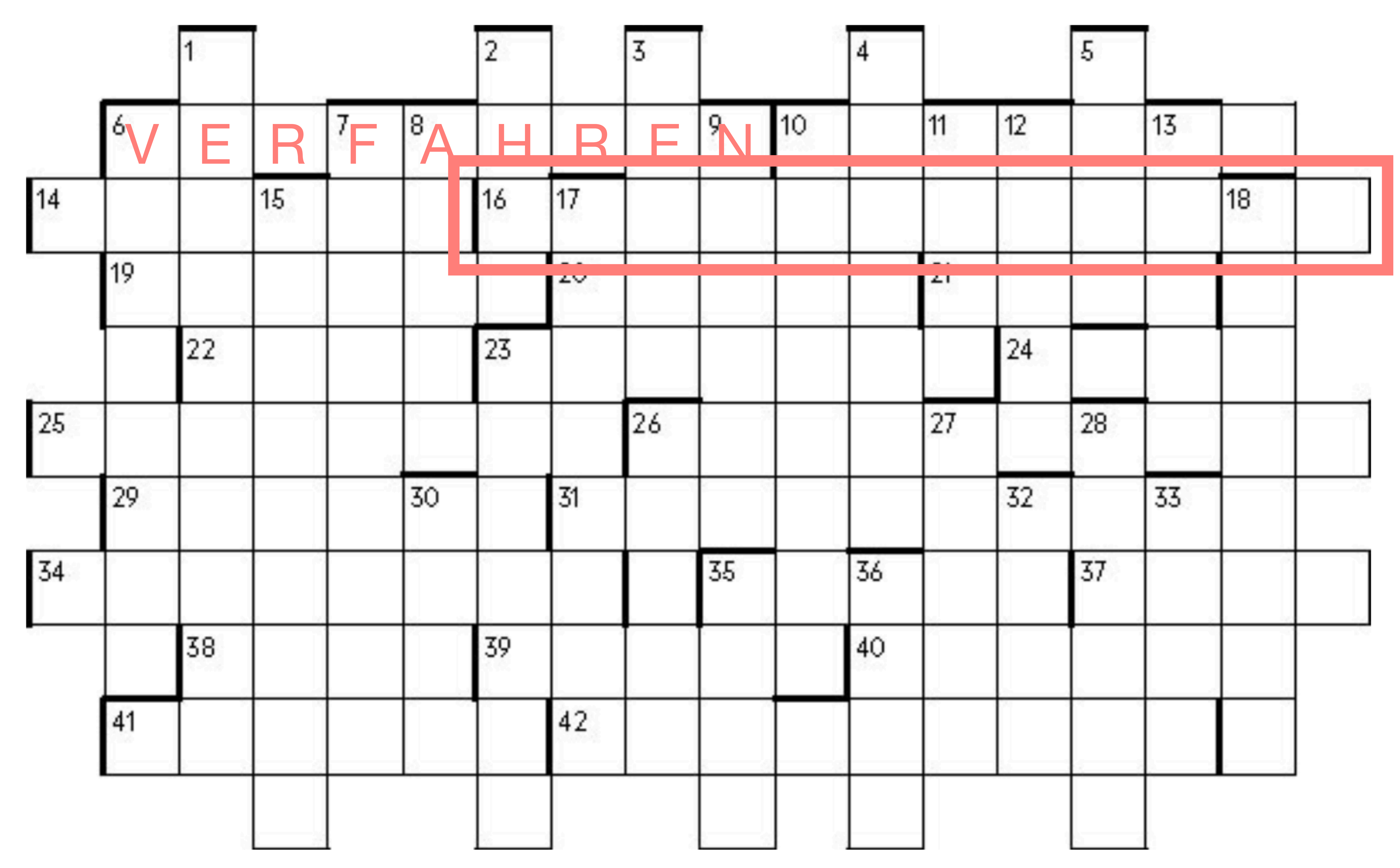
THINK OUTSIDE THE BOX RIDDLE

ZEIT MAGAZIN - UM DIE ECKE GEDACHT

16 HORIZONTAL

Das Anlegen wärmender Kleidung? Das Veranlassen einer Einnahmebuchung zugunsten der Bank!

(Translates to: To put on warming clothes? Causing the bank to draw money for their own profit!)



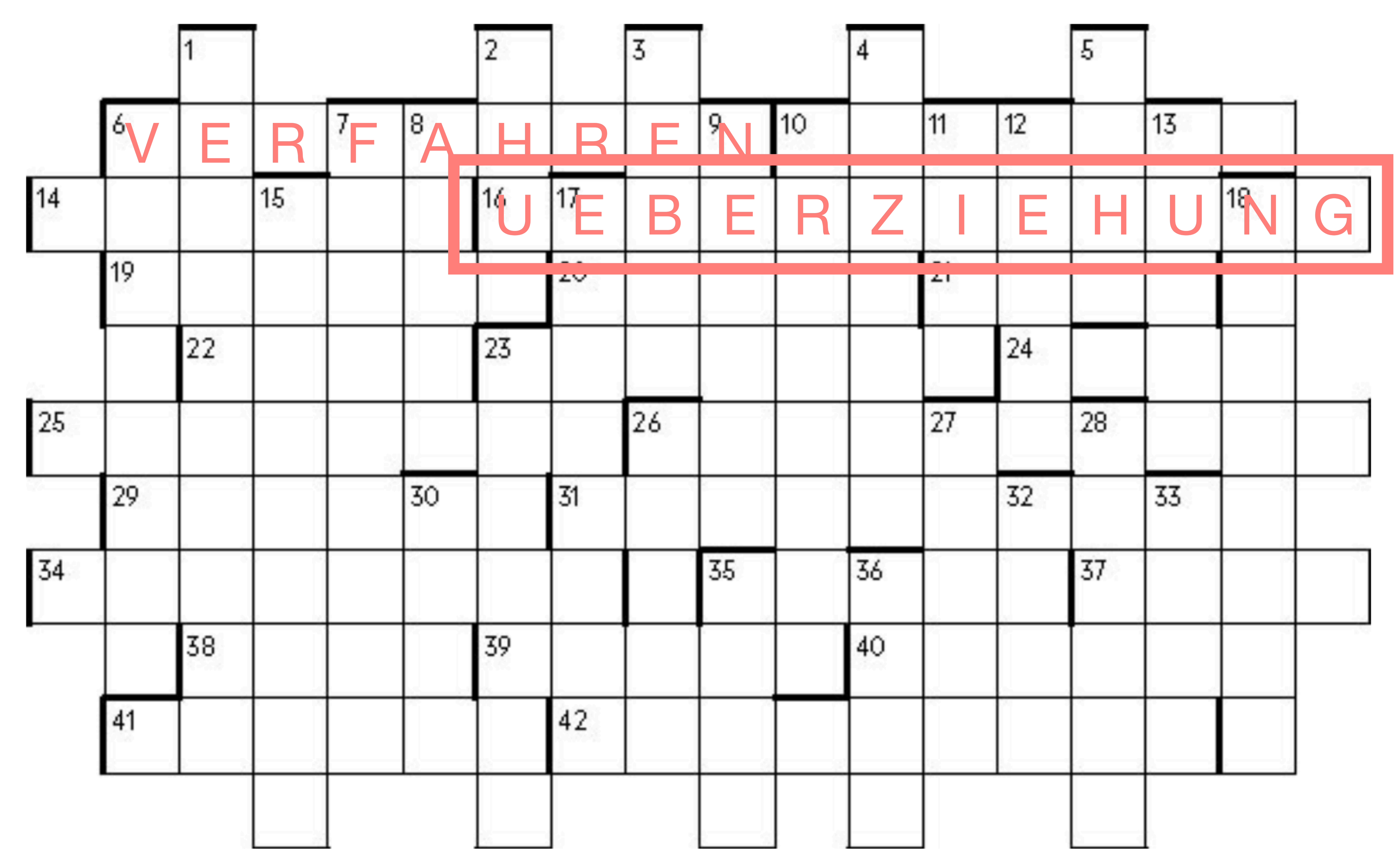
THINK OUTSIDE THE BOX RIDDLE

ZEIT MAGAZIN - UM DIE ECKE GEDACHT

ANSWER

Das Anlegen wärmender Kleidung? Das Veranlassen einer Einnahmebuchung zugunsten der Bank!
Überziehung

(Translates to: To put on warming clothes? Causing the bank to draw money for their own profit!
Overdrawing)



RULE 2

Generate 2 sentences: One ending on a question mark, one on an exclamation mark. The exclamation describes the meaning of the word we are looking for. The question describes the meaning of the same word but misinterpreted.

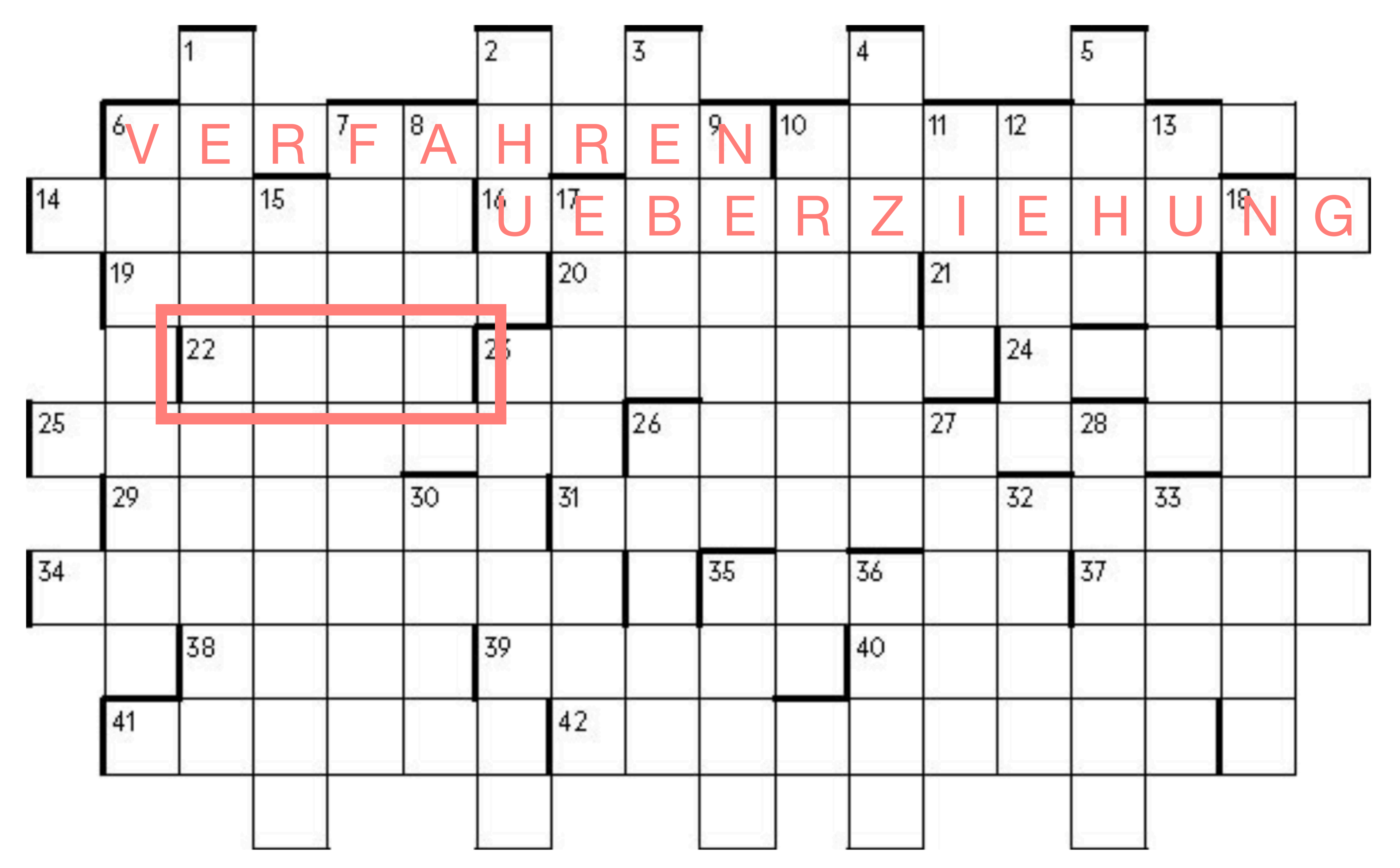
THINK OUTSIDE THE BOX RIDDLE

ZEIT MAGAZIN - UM DIE ECKE GEDACHT

22 HORIZONTAL

Der Ring macht …, und Ringe sind's, die eine Kette machen (Fr. Schiller)

(Translates to: The ring makes …, and it is rings that make a chain)



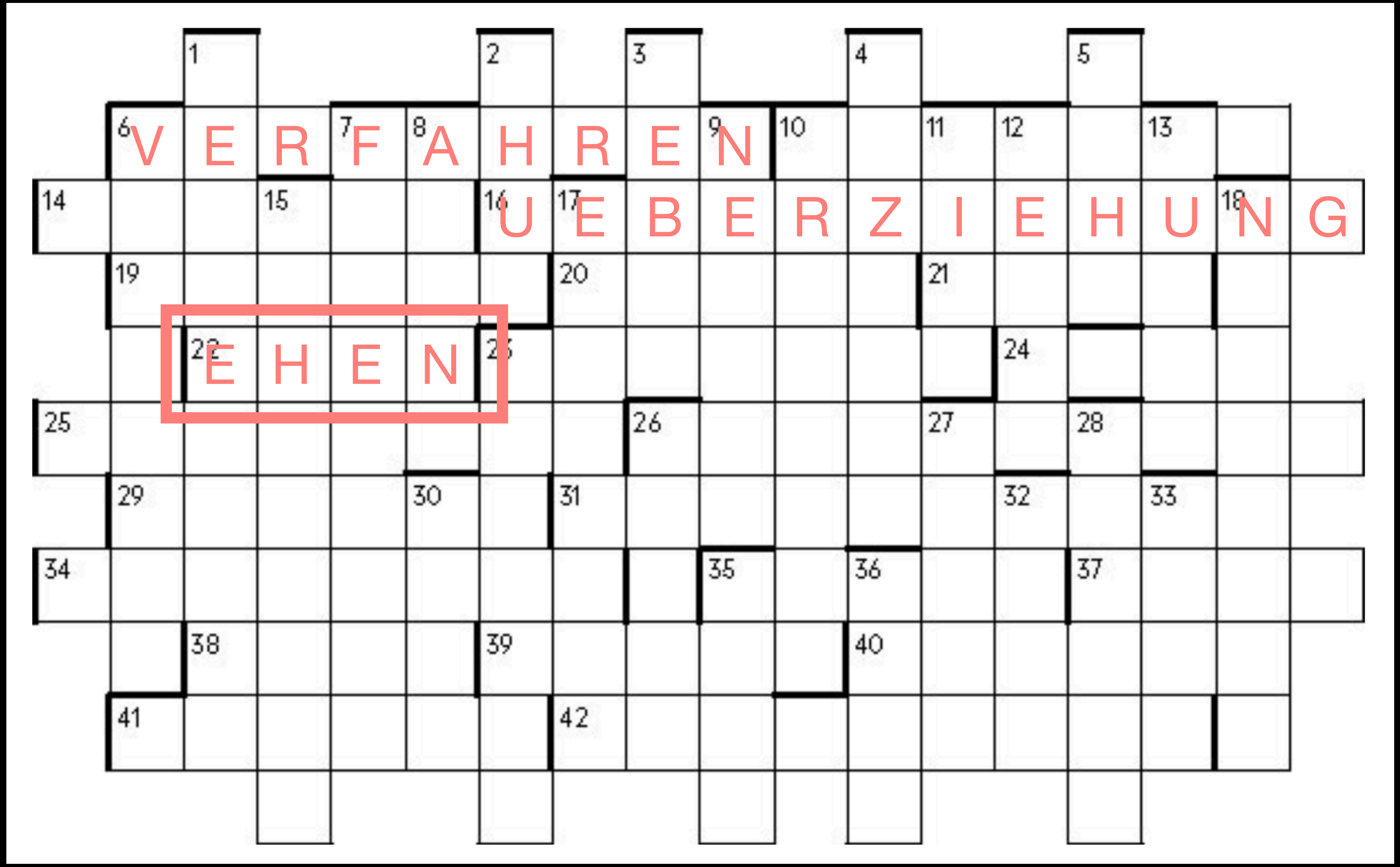
THINK OUTSIDE THE BOX RIDDLE

ZEIT MAGAZIN - UM DIE ECKE GEDACHT

ANSWER

Der Ring macht Ehen, und Ringe sind's, die eine Kette machen (Fr. Schiller)

(Translates to: The ring makes marriages, and it is rings that make a chain)



RULE 3

Use an existing citation, leave out a word.

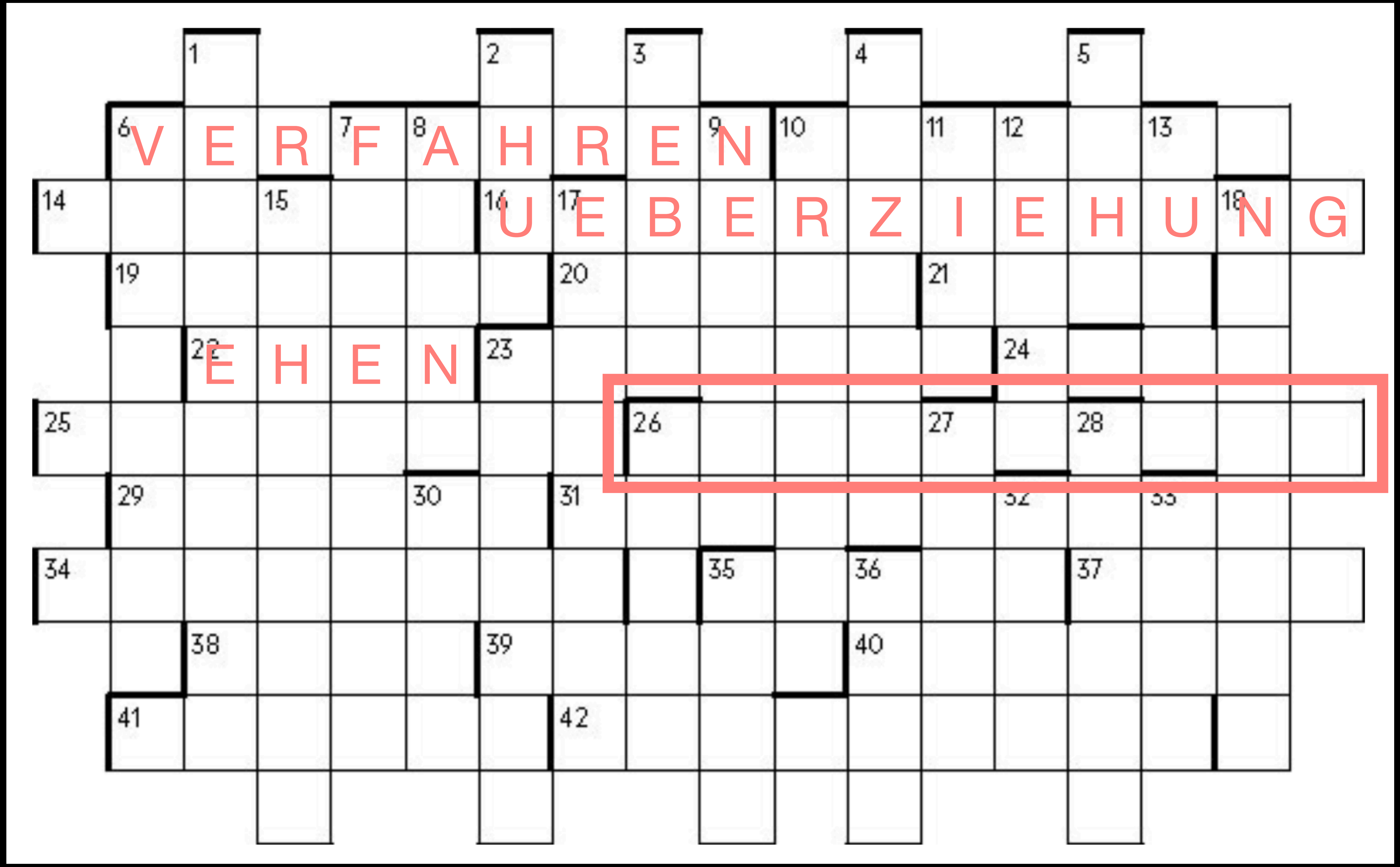
THINK OUTSIDE THE BOX RIDDLE

ZEIT MAGAZIN - UM DIE ECKE GEDACHT

26 HORIZONTAL

Rastplatz auf dem Weg von Hang zu Glas

(Translates to: Parking place on the way from hillside to glass)



THINK OUTSIDE THE BOX RIDDLE

ZEIT MAGAZIN - UM DIE ECKE GEDACHT

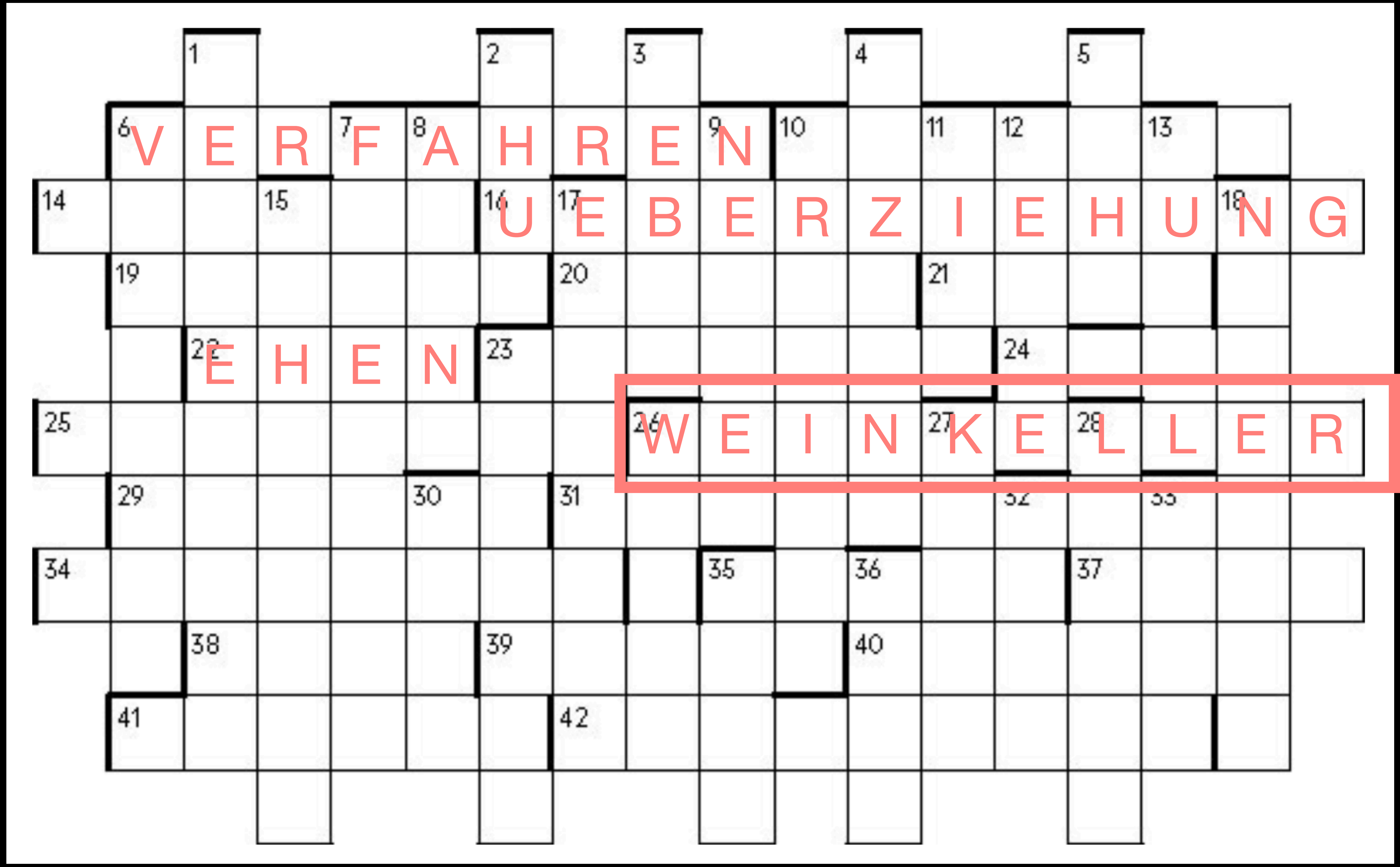
ANSWER

Rastplatz auf dem Weg von Hang zu Glas

Weinkeller

(Translates to: Parking place on the way from hillside to glas)

wine cellar



RULE 4

Shortly describe a process or concept using ambiguous words.

THANK YOU

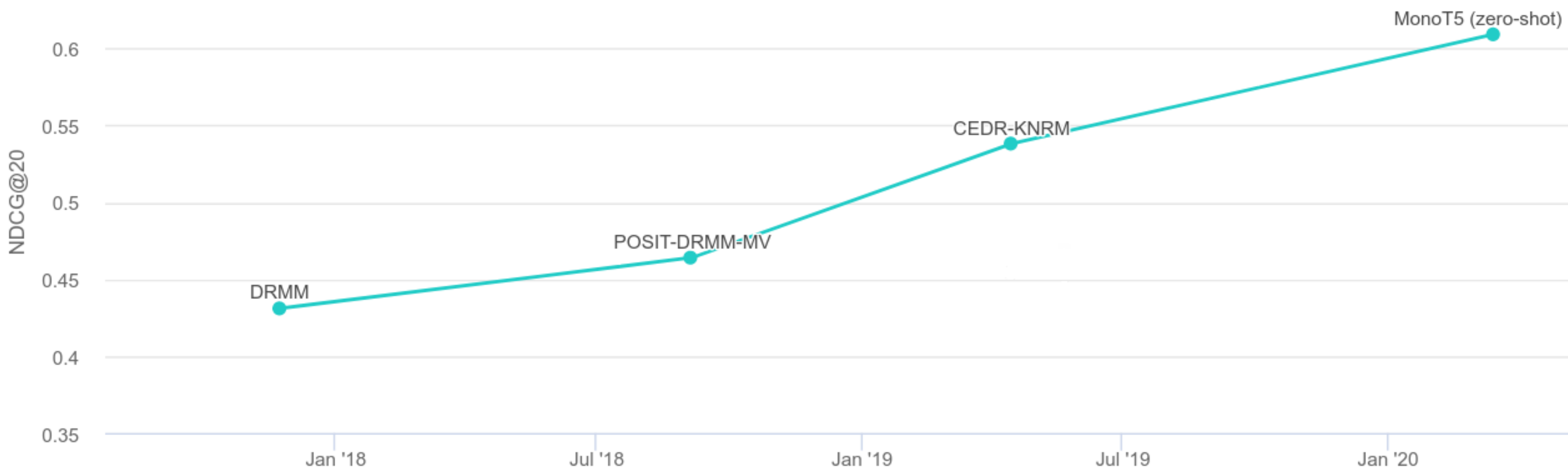
Zero Shot or Not?
The Effects of Train–Test Leakage on Neural Retrieval Models

Maik Fröbe, Christopher Akiki, Martin Potthast, Matthias Hagen

WHEN YOU TRAIN PREDICTIVE MODELS ON INPUT FROM YOUR USERS, IT CAN LEAK INFORMATION IN UNEXPECTED WAYS.

The Effects of Train–Test Leakage on Neural Retrieval Models

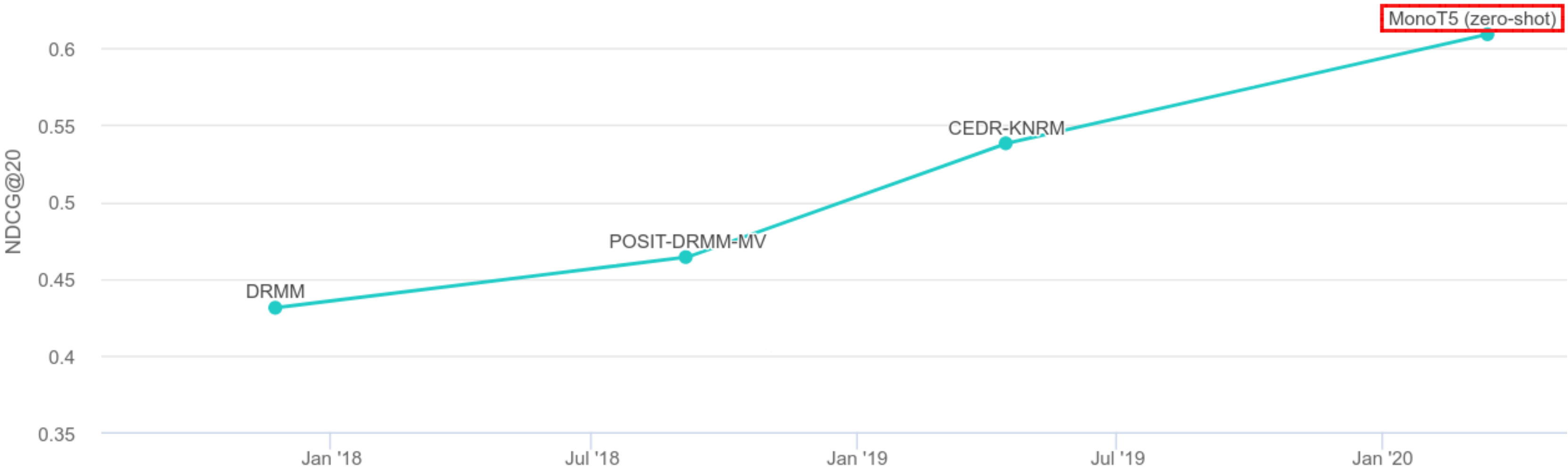
Motivation: Leaderboard for Retrieval Effectiveness on Robust04



- Robust04: 249 test queries with dense judgments
  - Traditional setup with cross-validation

The Effects of Train–Test Leakage on Neural Retrieval Models

Motivation: Leaderboard for Retrieval Effectiveness on Robust04

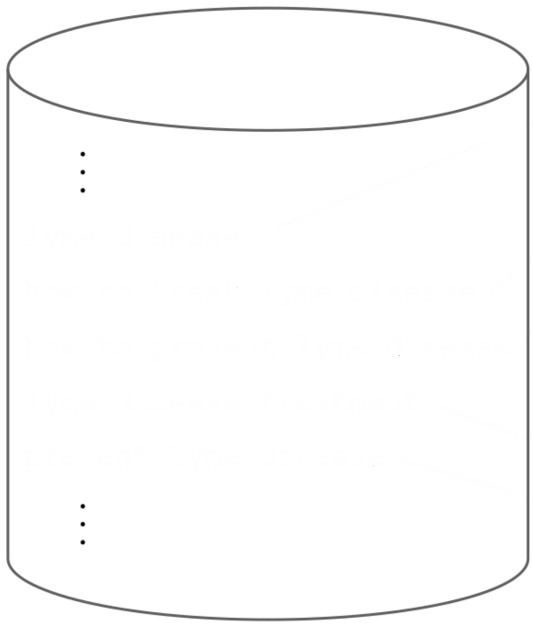


- Robust04: 249 test queries with dense judgments
  - Traditional setup with cross-validation
- MonoT5 (zero-shot)
  - Trained only on MS MARCO (> 10 million queries available)

The Effects of Train–Test Leakage on Neural Retrieval Models

Overlapping Queries for Topic 441 of Robust04

MS MARCO



Robust04

Title: lyme disease

Description: How do you prevent and treat Lyme disease?

Narrative: Documents that discuss current prevention and treatment techniques for Lyme disease are relevant. Reports of research on new treatments of the disease are also relevant.

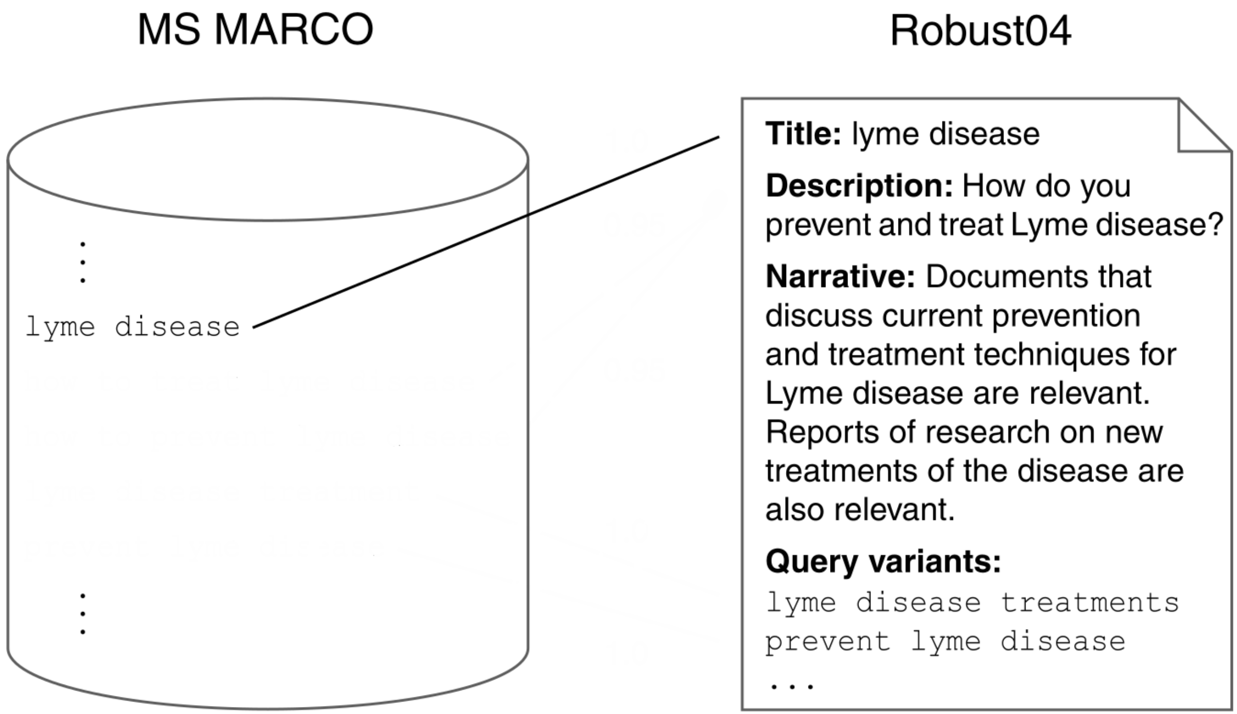
Query variants:

```
lyme disease treatments
prevent lyme disease
...
```

- Train on many queries
- Test on 249 queries

The Effects of Train–Test Leakage on Neural Retrieval Models

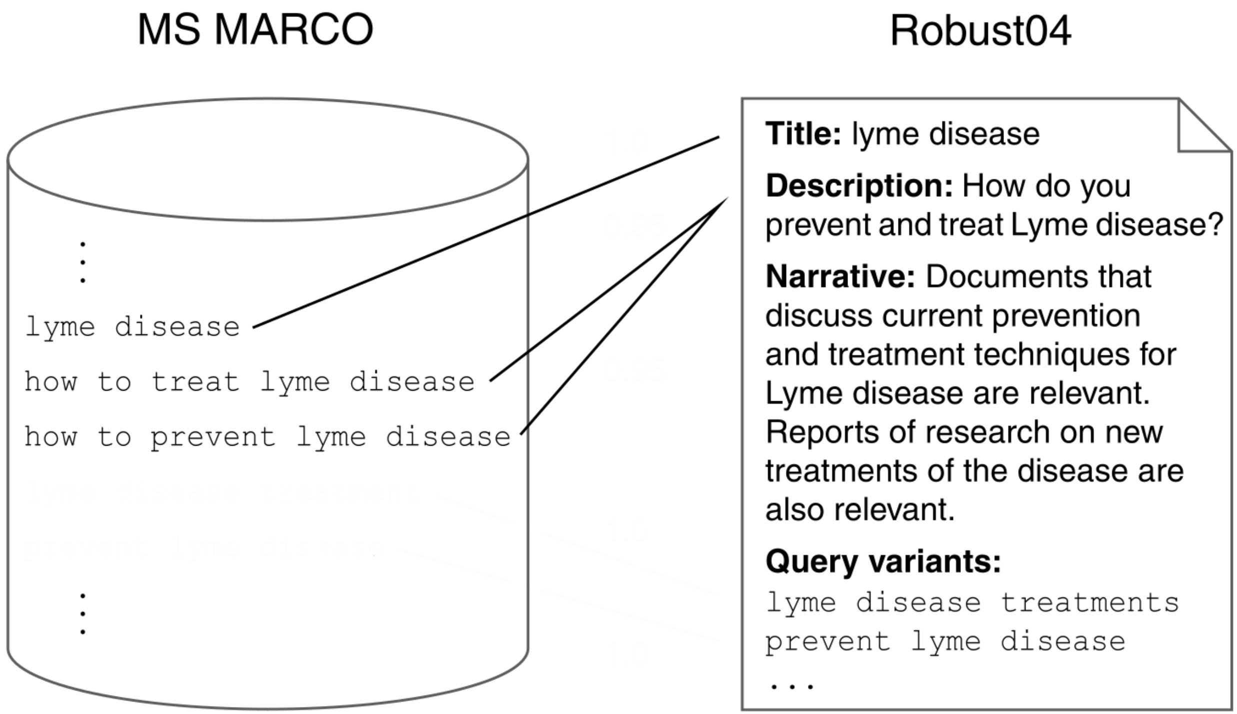
Overlapping Queries for Topic 441 of Robust04



- Train on many queries
- Test on 249 queries
- Nearest-neighbor search for overlapping queries (S-BERT + FAISS)

The Effects of Train–Test Leakage on Neural Retrieval Models

Overlapping Queries for Topic 441 of Robust04

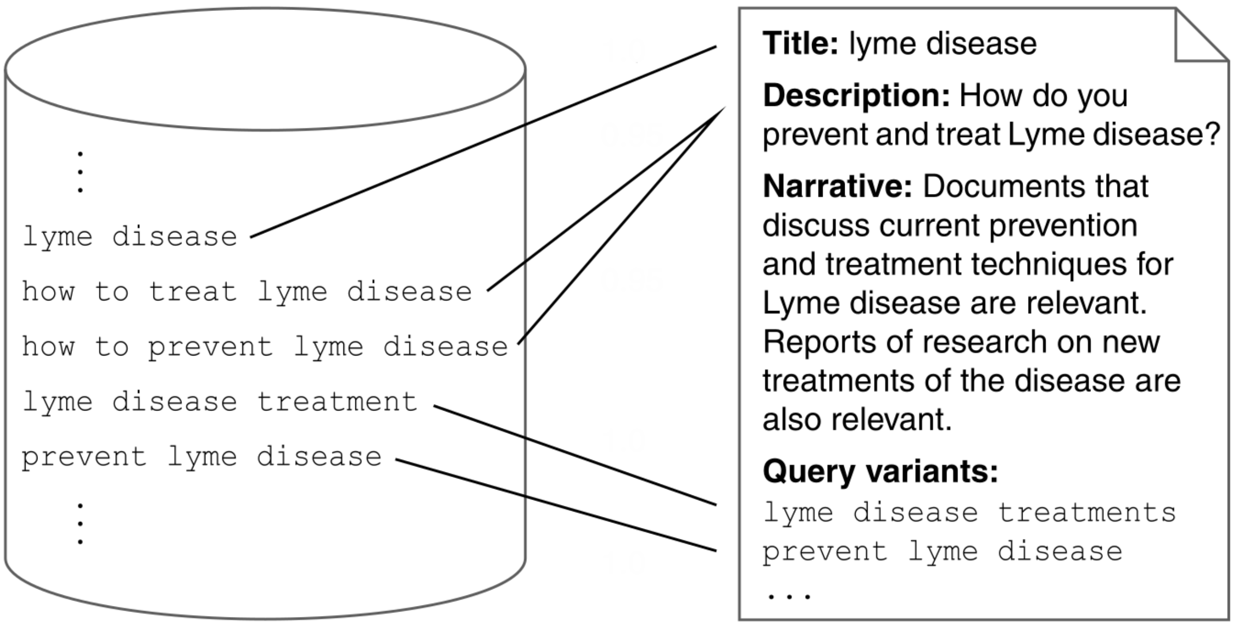


- Train on many queries
- Test on 249 queries
- Nearest-neighbor search for overlapping queries (S-BERT + FAISS)

Overlapping Queries for Topic 441 of Robust04

MS MARCO

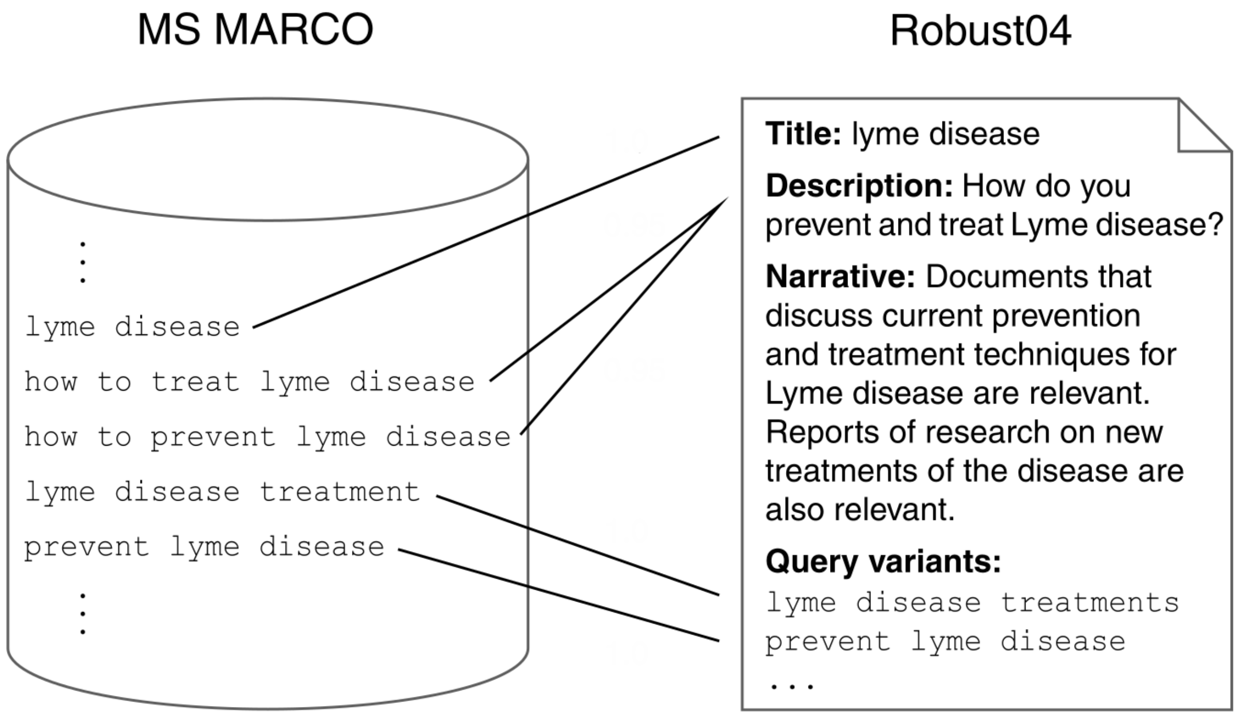
Robust04



- Train on many queries
- Test on 249 queries
- Nearest-neighbor search for overlapping queries (S-BERT + FAISS)

The Effects of Train–Test Leakage on Neural Retrieval Models

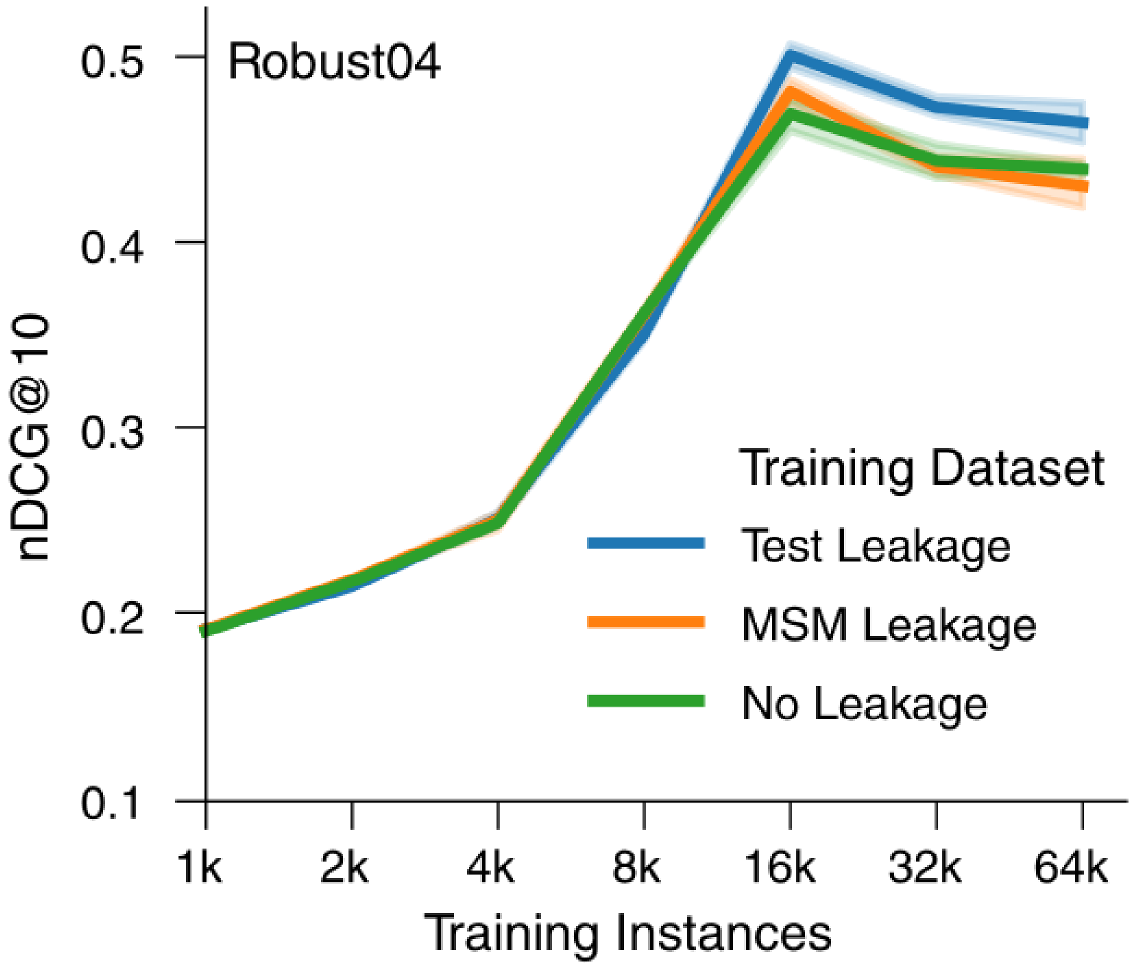
Overlapping Queries for Topic 441 of Robust04



- Train on many queries
- Test on 249 queries
- Nearest-neighbor search for overlapping queries (S-BERT + FAISS)
- 172 of 249 test queries from Robust04 occur in MS MARCO (69%)
- Is the evaluation of MonoT5 invalidated by overlapping queries?

The Effects of Train–Test Leakage on Neural Retrieval Models

Effectiveness of Retrieval Models



The Effects of Train–Test Leakage on Neural Retrieval Models

Takeaways

- Possible train–test leakage for models trained on MS MARCO
  - Potential to invalidate experiments
  - Default in PyTerrier/Pyserini/PyGaggle often trained on MS MARCO
  - Only few training instances overlap: Impact measurable, but negligible
- Future work:
  - Effects on Dense Retrieval models
  - Practical consequences for real search engines

The Effects of Train–Test Leakage on Neural Retrieval Models

Takeaways

- Possible train–test leakage for models trained on MS MARCO
  - Potential to invalidate experiments
  - Default in PyTerrier/Pyserini/PyGaggle often trained on MS MARCO
  - Only few training instances overlap: Impact measurable, but negligible
- Future work:
  - Effects on Dense Retrieval models
  - Practical consequences for real search engines

Thank You!

Names
(just a very short snippet of an overview)

Webis Flash Talks 2022
Sebastian Günther

Names… really?

Specifically: author names

- There is not *"the one way"* to spell a name
  - Umlauts, accent marks
  - Abbreviations
  - Dashes

- Surname variety:
  - Germany: ~52627[1] (including foreign names!)
  - Han Chinese: ~2000 (50% use 19, 87% use 100)[2]

[1]https://www.namenforschung.net/en/dfd/dictionary/list/; [2]https://en.wikipedia.org/wiki/Chinese_surname

Varianten [Bearbeiten | Quelltext bearbeiten]

Der Name Meier gehört zu den häufigsten Familiennamen im deutschen Sprachraum.

Allein zur Schreibweise *Meyer* gab es in den Telefonbüchern 2005 in Deutschland über 100.000 Einträge (Platz 5), mit allen Varianten waren es ungefähr 260.000:

- Meyer (100.638)
- Meier (44.996)
- Maier (40.246)
- Mayer (39.217)
- Mayr (8.105)
- Mair (2.665)
- Meir (1.061)
- Majer (885)
- Meyr (653)
- Meijer (191)
- Mejer (49)
- Maiers (30)
- Maijer (13)
- Maihr (7)
- Maiyer (2)
- Meiyer (2)
- Myer
- Mayers
- Meiering
- Meyering
- Meiers
- Meyers
- Myers
- Mayring
- Meyrink
- Meyerson

Stand: Herbst 2002 (Meijer, Maijer: Winter 2005)

26!

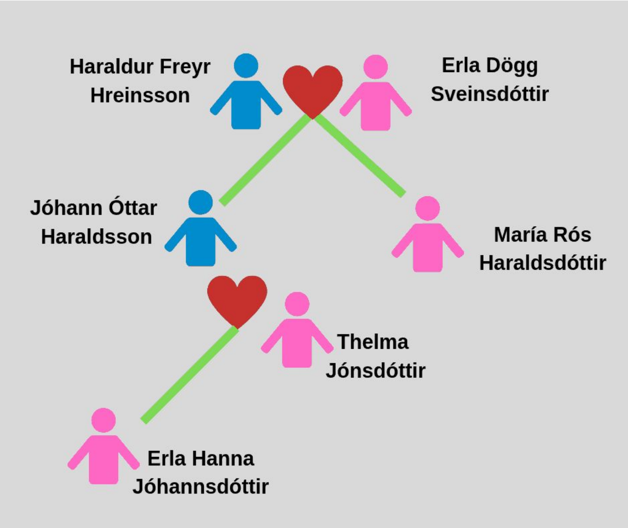
Hello /r/unexpectedfactorial/

https://de.wikipedia.org/wiki/Meier_(Familienname)

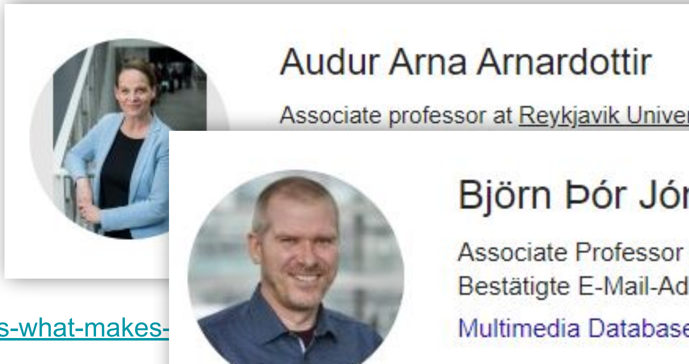
Names on "Easy Level"

- In "our" western world
- No titles (e.g., "Sir", "Dr.", "Prof.", etc.)

- Germany:
  - Up to 5 first names[1] (depends)
  - Double first names are okay: "Karl-Heinz" or "Karl Heinz"
  - Hyphenated double surname (marriage)
- Iceland: suffix "son" (male) or "dóttir" (female)
  - What about "Auður"?
  - Also: names may change depending on the context[2]



2



[1]https://www.rosepartner.de/namensrecht-name.html; [2]https://www.icelandtravel.is/blog/icelandic-names-what-makes-

Author Names According to APA (easy)

- Karl-Heinz Meier-Maier → Meier-Maier, K.-H.
- Karl Heinz Meier-Maier → Meier-Maier, K. H.

- Karl Meyer Sr → Meyer, K., Sr.

- Johannes Diderik van der Waals → van der Waals, J. D.

- …

https://libguides.singaporetech.edu.sg/citation/author

Theory as Liberatory Practice

bell hooks†

Let me begin by saying that I came to theory because I was hurting—the pain within me was so intense that I could not go on living. I came to theory desperate, wanting to comprehend—to grasp what was happening around and within me. Most importantly, I wanted to make the hurt go away. I saw in theory then a location for healing.

I came to theory young, when I was still a child. In *The Significance of Theory* Terry Eagleton says:

> Children make the best theorists, since they have not yet been educated

hooks, bell. Theory as liberatory practice. Yale JL & Feminism, 1991, Vol. 4 , p1.

https://www.reddit.com/r/funny/comments/80csc6/complete_chinese_keyboard/

Back to Chinese Names (a.k.a “not so easy”)

- 劉 / 柳 / 留 / 六 - Liu
- Romanization systems:
  - Pinyin (chinese)
  - Rōmaji (japanese)

General differences:

- Chinese names start with their “last name”
- Double names: 小平 → Xiaoping / Xiao-Ping / Xiao Ping

Award Recipient: ChengXiang Zhai (University of Illinois at Urbana-Champaign)

https://sigir.org/sigir2021/awards/

- They address each other with their full name
- Some use a westernized version of their name
  - Zhang Chen → James Chen Zhang

https://culturalatlas.sbs.com.au/chinese-culture/chinese-culture-naming

Non-Western Names in General (a.k.a impossible)

- Indonesia: some people only have **one name**
- India: names can include a **family, village, middle, or pet name**
  - and their order may also vary
- Muslim names:
  - **May include** a family name
  - **May not include** a family name
- Myanmar:
  - People usually only have given names
  - Consist of two to three syllables
  - Example: Aung San Suu Kyi (but: includes the father's name: Aung San)

https://libguides.singaporetech.edu.sg/citation/author

Non-western Names

APA examples:

- Muslim names don’t change unless specified otherwise
    - Sulaiman Abdullah Saif Al Nasser → Sulaiman Abdullah Saif Al Nasser

- bin/binti = “son/daughter of”
    - Mahathir bin Muhammed → Mahathir Muhammed

Conclusion

- In general: author names are systematically derived from peoples' real names

(…most of the time)

- Only works for western languages
- There are various exceptions, especially in non-western schemes
- You can decide to have you name spelled in a special way

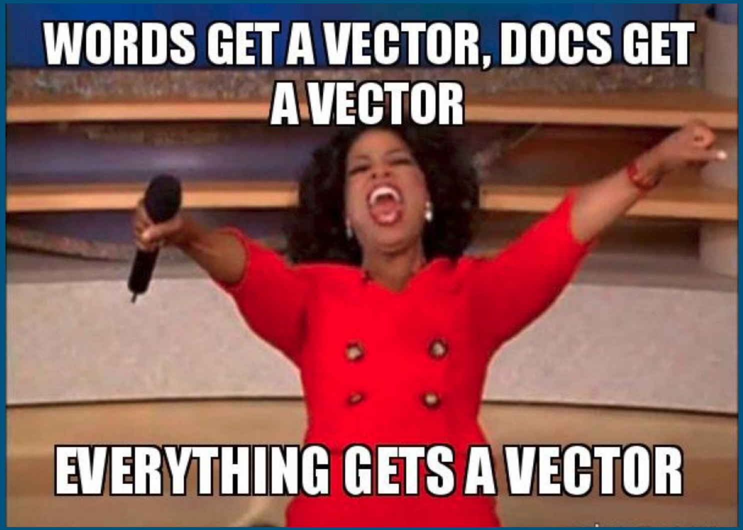
- Rules? Suggestions!

Thank you!

Evaluation of word embeddings

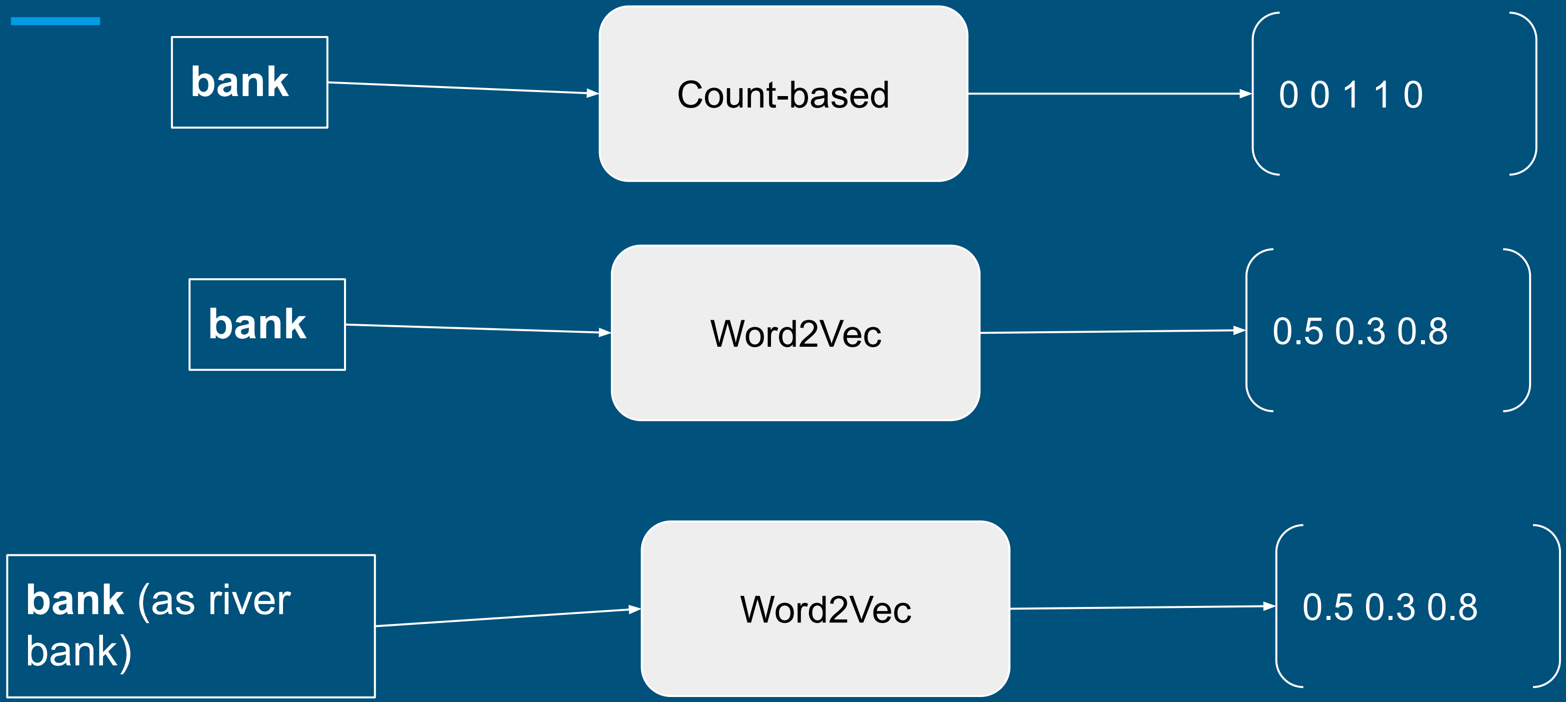
Master Thesis
Hannes Hansen

Word embeddings

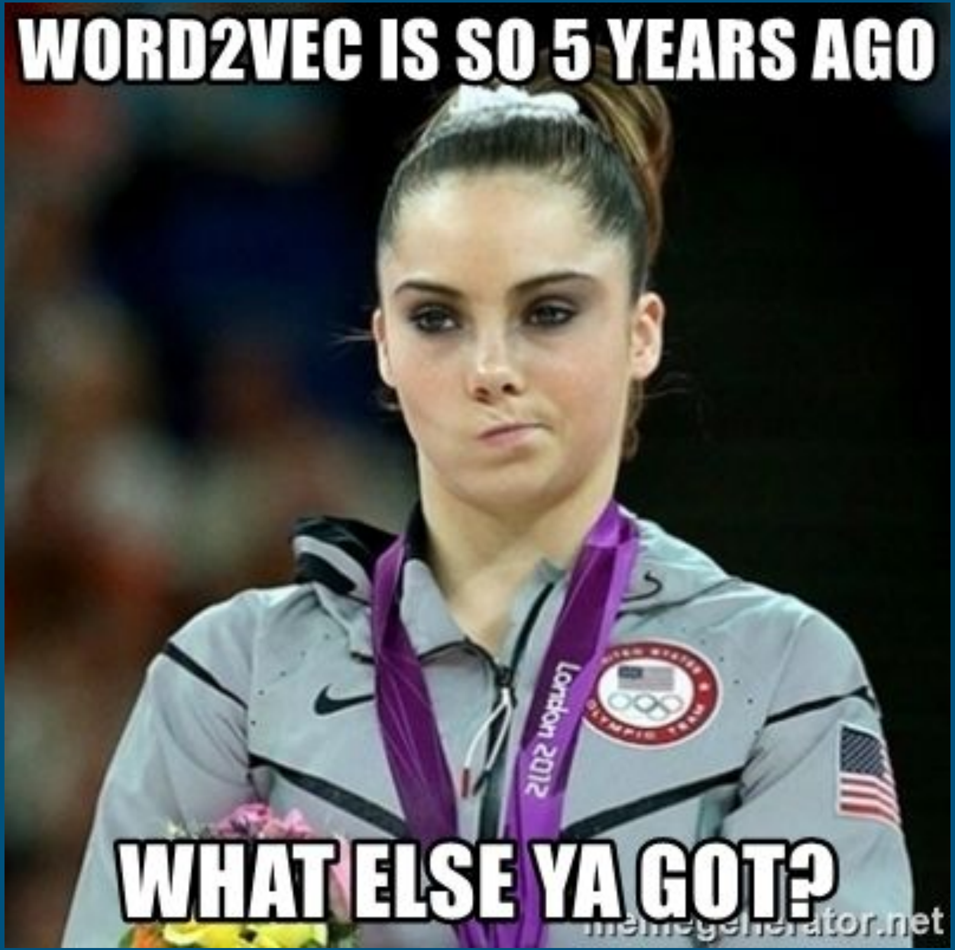


Foundation of a lot of NLP downstream tasks!

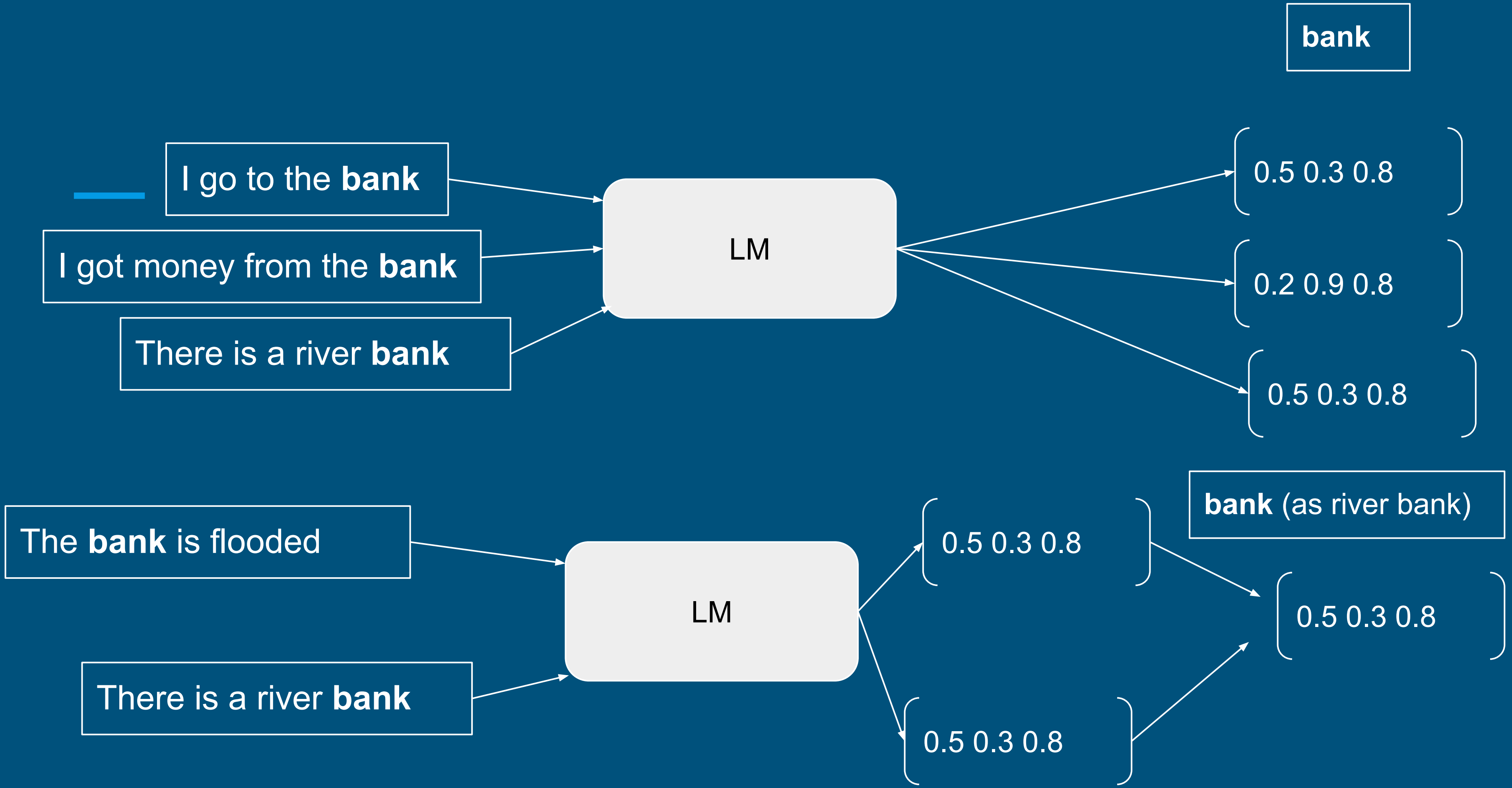
Recent work



bank
Count-based
0 0 1 1 0
bank
Word2Vec
0.5 0.3 0.8
bank (as river bank)
Word2Vec
0.5 0.3 0.8


WORD2VEC IS SO 5 YEARS AGO
London 2012
WHAT ELSE YA GOT?

Leverage the power of large pretrained language models to generate word embeddings


bank
I go to the **bank**
I got money from the **bank**
There is a river **bank**
LM
0.5 0.3 0.8
0.2 0.9 0.8
0.5 0.3 0.8
The **bank** is flooded
There is a river **bank**
LM
0.5 0.3 0.8
0.5 0.3 0.8
bank (as river bank)
0.5 0.3 0.8

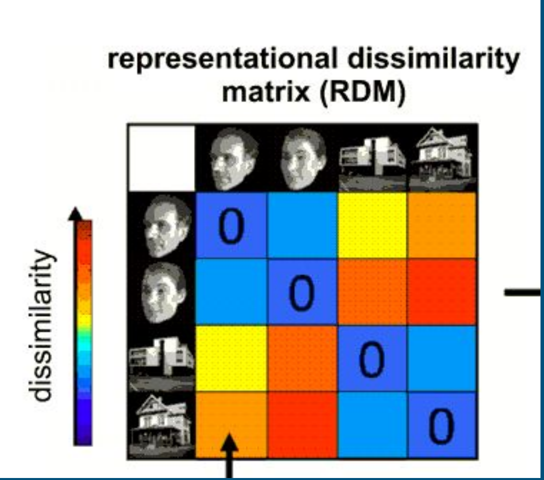
Goal

1. Static embeddings (fast and usable)
2. Using power of large language models
3. multiple levels (word, word senses, concepts)

Pipeline

1. Annotate corpus with senses
2. Extract contextualized word embeddings from LLMs
3. Aggregate to a static embedding per word/word sense
4. Evaluate

Evaluation



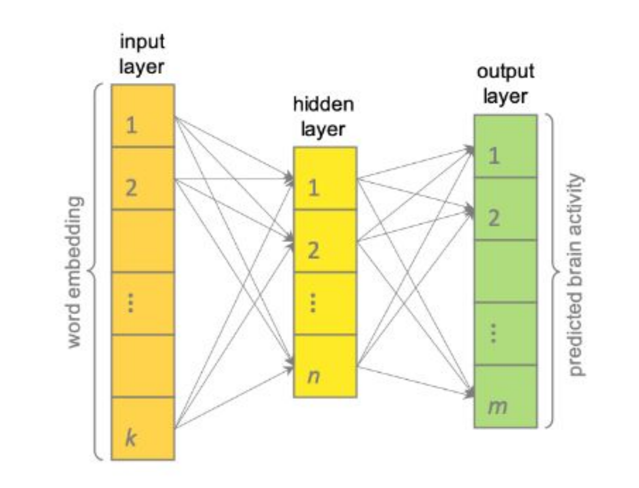


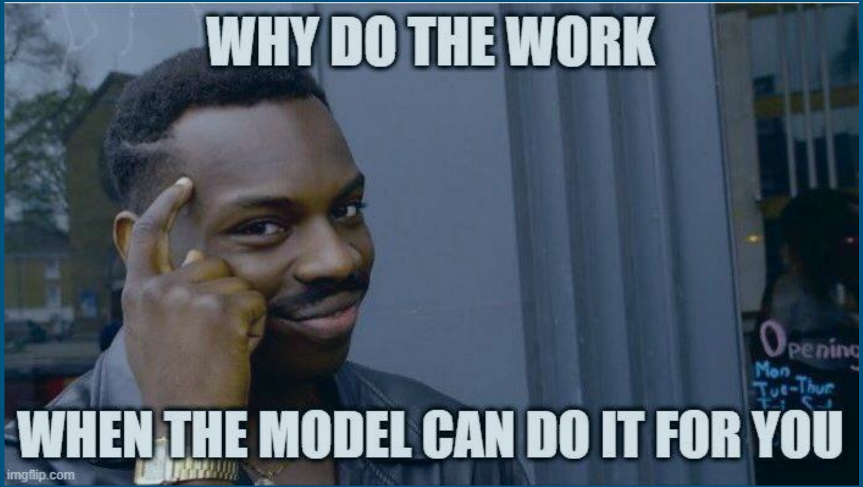
Figure 2: Neural architecture of regression models.

Low Context Word Prediction

Students: Ronald Kurniawan & Abdul Hanan Khan
Supervisors: Matti Wiegman & Michael Völske
Bauhaus Universität Weimar

Goal

- Masked Language Modelling / Fill in the blanks
- Low context → 3 – 9 words per query
- Single mask → “Never gonna ? you up”

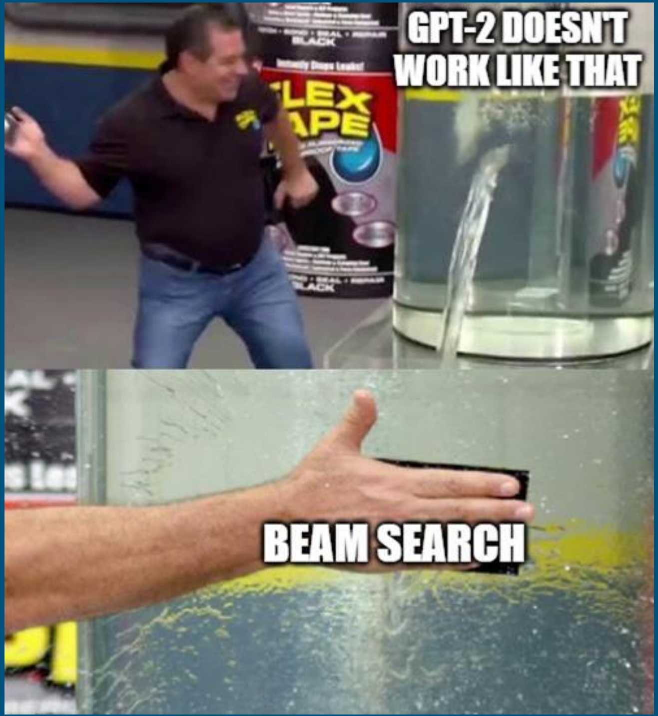


Approach: BERT

- Primarily trained for two purposes:
  - Masked language modeling
  - Next sentence prediction
- It is bidirectional (left to right as well as right to left)
- Initially trained on high context data, therefore fine tuning is required.

Approach: GPT-2

- Doesn't work out-of-the-box, need to be modified
- Beam search approach
- 2 functions:
    - Predict next word →
    - Recalculate probabilities →



Approach: GPT-2

Query: “I work as a ? at the restaurant”

“I work as a”

Approach: GPT-2

Query: “I work as a ? at the restaurant”

“I work as a” ⟶

“I work as a writer”
“I work as a freelance”
“I work as a consultant”
…
“I work as a developer”

Predict next word

Approach: GPT-2

Query: “I work as a ? at the restaurant”

“I work as a” → “I work as a writer” / “I work as a freelance” / “I work as a consultant” / … / “I work as a developer” → “I work as a chef at the restaurant” / “I work as a sales at the restaurant” / “I work as a manager at the restaurant” / … / “I work as a professor at the restaurant”

		“I work as a writer”		“I work as a chef at the restaurant”
		“I work as a freelance”		“I work as a sales at the restaurant”
“I work as a”	→	“I work as a consultant”	→	“I work as a manager at the restaurant”
		…		…
		“I work as a developer”		“I work as a professor at the restaurant”
	Predict next word		Recalculate probabilities	

Approach: GPT-2

Query: “I work as a ? at the restaurant”

“I work as a” → “I work as a writer”
“I work as a freelance”
“I work as a consultant”
…
“I work as a developer”

→ “I work as a chef at the restaurant”
“I work as a sales at the restaurant”
“I work as a manager at the restaurant”
…
“I work as a professor at the restaurant”

Predict next word

Recalculate probabilities

Examples

Query: “? works as a nurse”

BERT results :

“She works as a nurse”
“He works as a nurse”
“Kate works as a nurse”

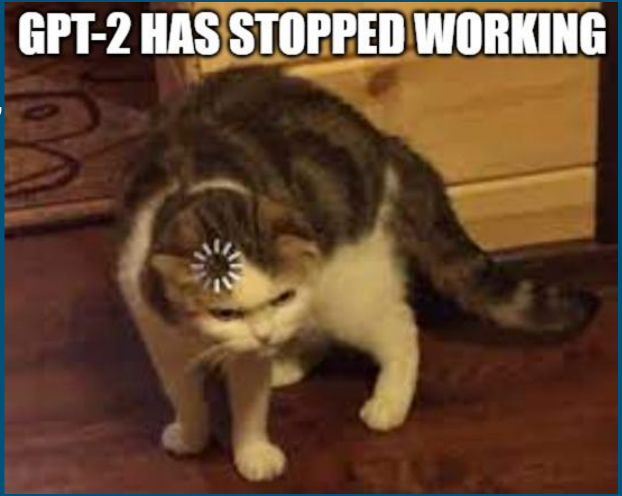
GPT-2 results: -

Query: “? works as a carpenter”

BERT results :

“He works as a carpenter”
“She works as a carpenter”
“John works as a carpenter”

GPT-2 results: -



Netspeak results:

? works as a nurse		
who works as a nurse	953	53%
she works as a nurse	842	47%

? works as a carpenter		
he works as a carpenter	314	39%
who works as a carpenter	228	29%
and works as a carpenter	169	21%
now works as a carpenter	84	11%

Examples

Query: “To be or not to ?”

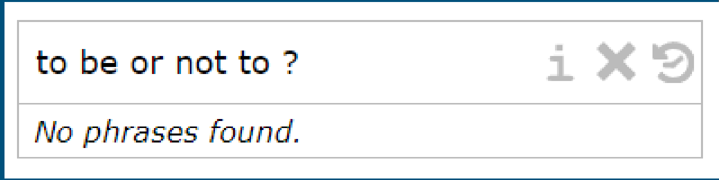
BERT results :

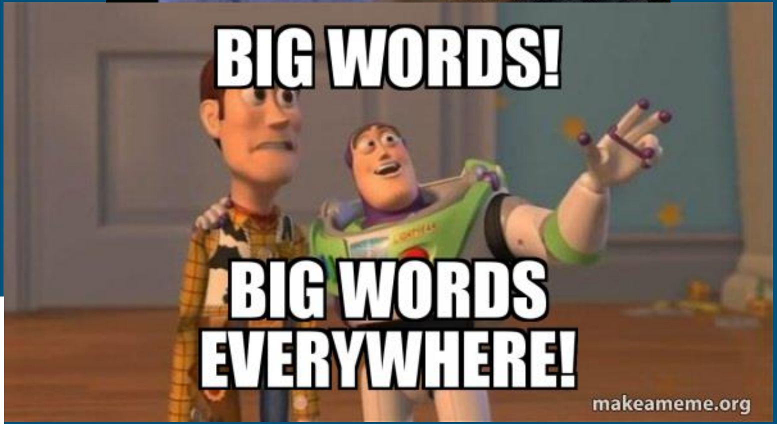
“To be or not to be”
“To be or not to do”
“To be or not to have”

GPT-2 results:

“To be or not to be”
“To be or not to have”
“To be or not to become

Netspeak results:





Well not really, but you get what we mean

Examples

Query: “BERT is ?”

BERT results :

“BERT is a”
“BERT is the”
“BERT is also”

GPT-2 results:

“BERT is a”
“BERT is the”
“BERT is not”

Netspeak results:

bert is ?		
bert is evil	5,500	100%

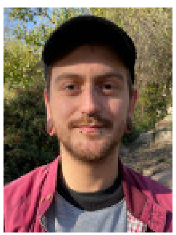
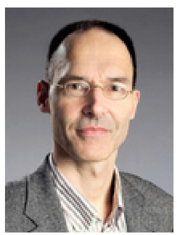
thank you ? much

thank you very much	2,600,000	64%
thank you so much	1,400,000	35%
thank you sooo much	15,000	0.4%
thank you soooo much	12,000	0.3%
thank you soo much	10,000	0.3%
thank you verry much	1,400	0.0%

Do Images Say More Than 1000 Arguments?

Webis Flash Talks 2022

Based on our ArgMining 2021 paper
"Image Retrieval for Arguments Using Stance-Aware Query Expansion"

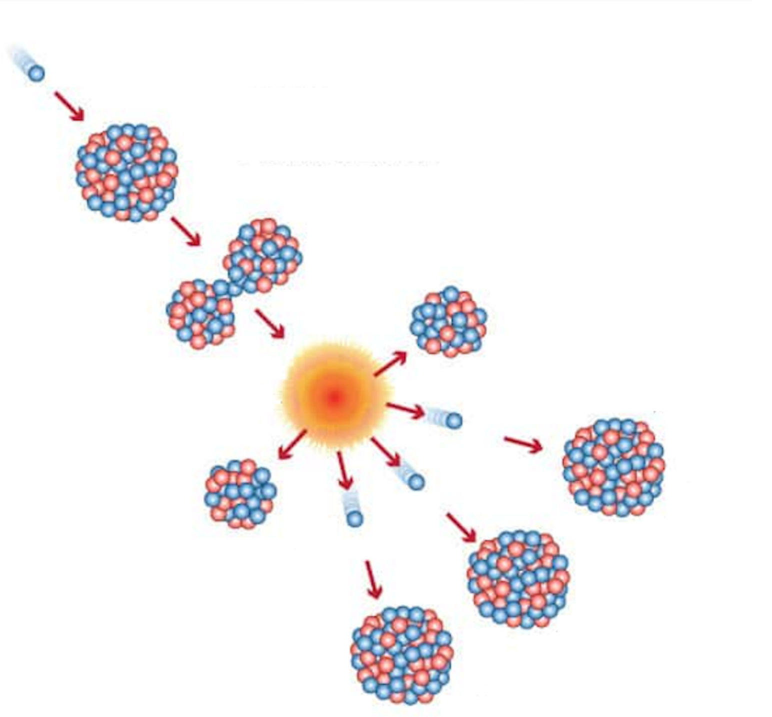

Johannes Kiesel[1]	Nico Reichenbach[2]	Benno Stein[1]	Martin Potthast[2]

Bauhaus-Universität Weimar [1]
UNIVERSITÄT LEIPZIG [2]
Webis [1,2]

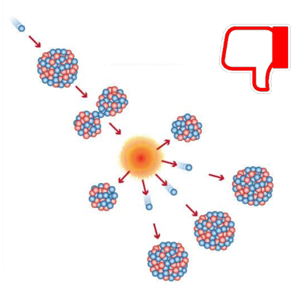
Do Images Say More Than 1000 Arguments?

Do Images Say More Than 1000 Arguments?

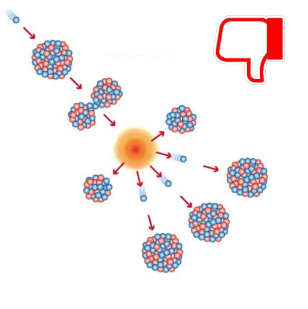
Do Images Say More Than 1000 Arguments?

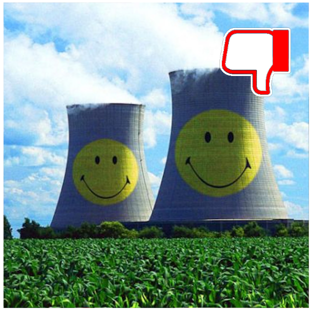


Do Images Say More Than 1000 Arguments?

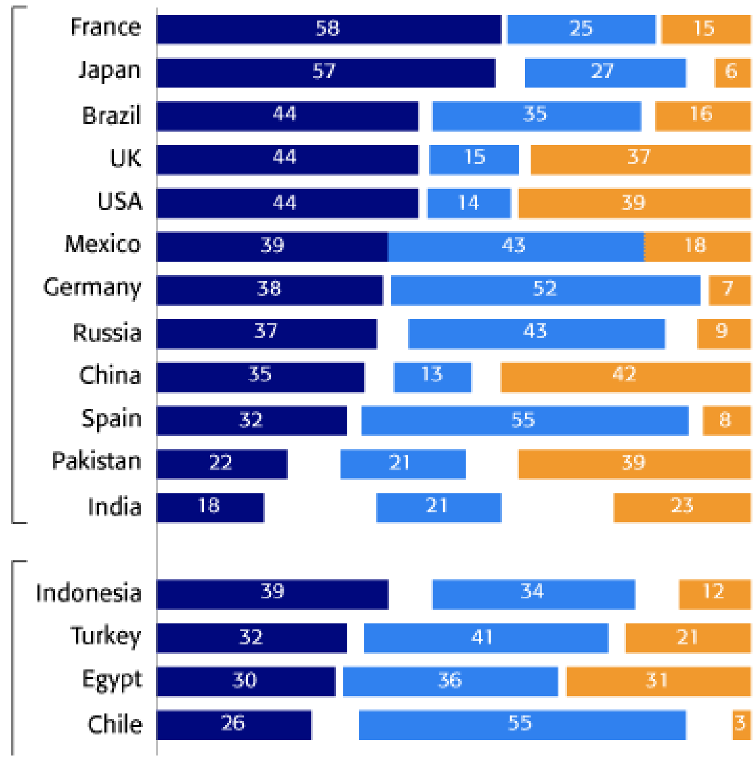


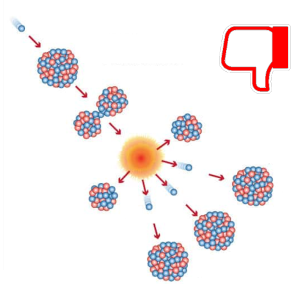
Do Images Say More Than 1000 Arguments?

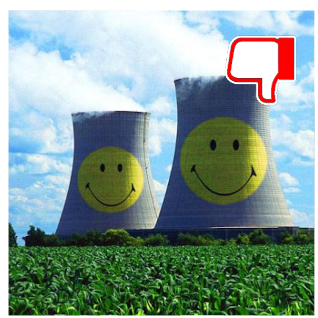




Do Images Say More Than 1000 Arguments?

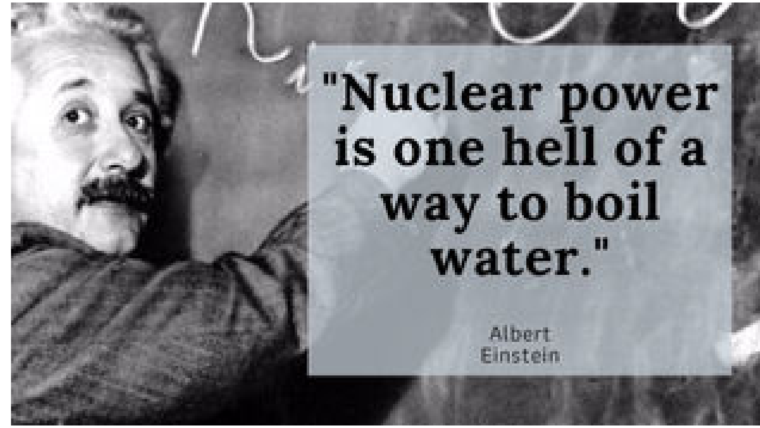


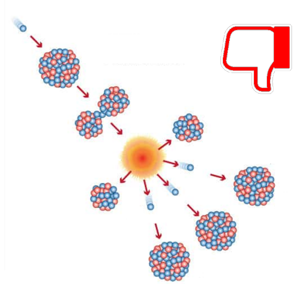


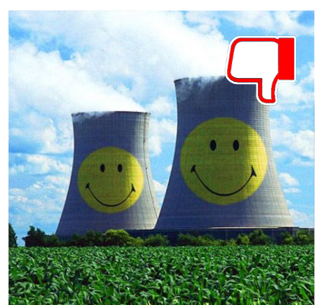


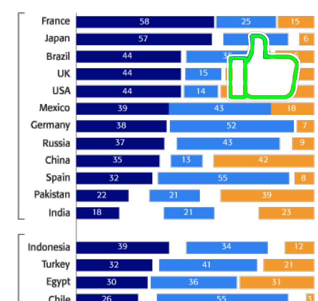
@KieselJohannes

Do Images Say More Than 1000 Arguments?







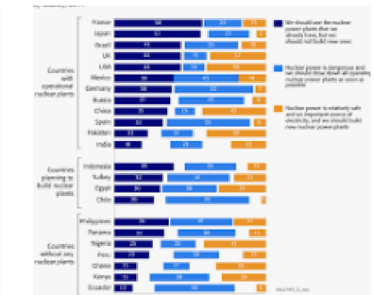


Nuclear energy

- [x] cartoons
- [x] covers
- [x] photos
- [x] posters
- [x] memes
- [x] protesters
- [x] quotes
- [x] statistics

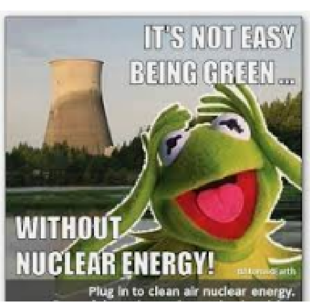
Can Nuclear power save us? – NUES
web.northeastern.edu

Californians for Green Nuclear Power
cgnp.org



Nuclear Energy Grows: Global Poll ...
globescan.com

marketers oppose clean energy subsidies ...
redgreenandblue.org



Orano U.S. on | Nuclear energy, ...
pinterest.dk

zero emissions must use nuclear energy
policyoptions.irpp.org

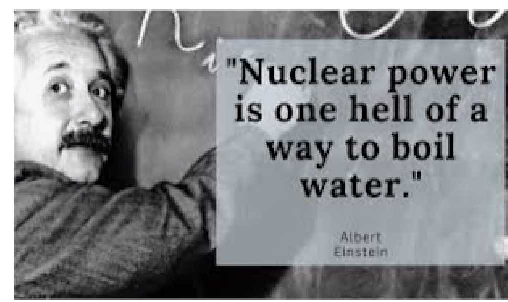
Anti-nuclear energy protestors take ...
alamy.com

Business Cartoon | TOONPOOL
toonpool.com

Anti-Nuclear Opposition | comparenuclear
comparenuclear.wordpress.com

No to nuclear: Japan wants reactors ...
aljazeera.com

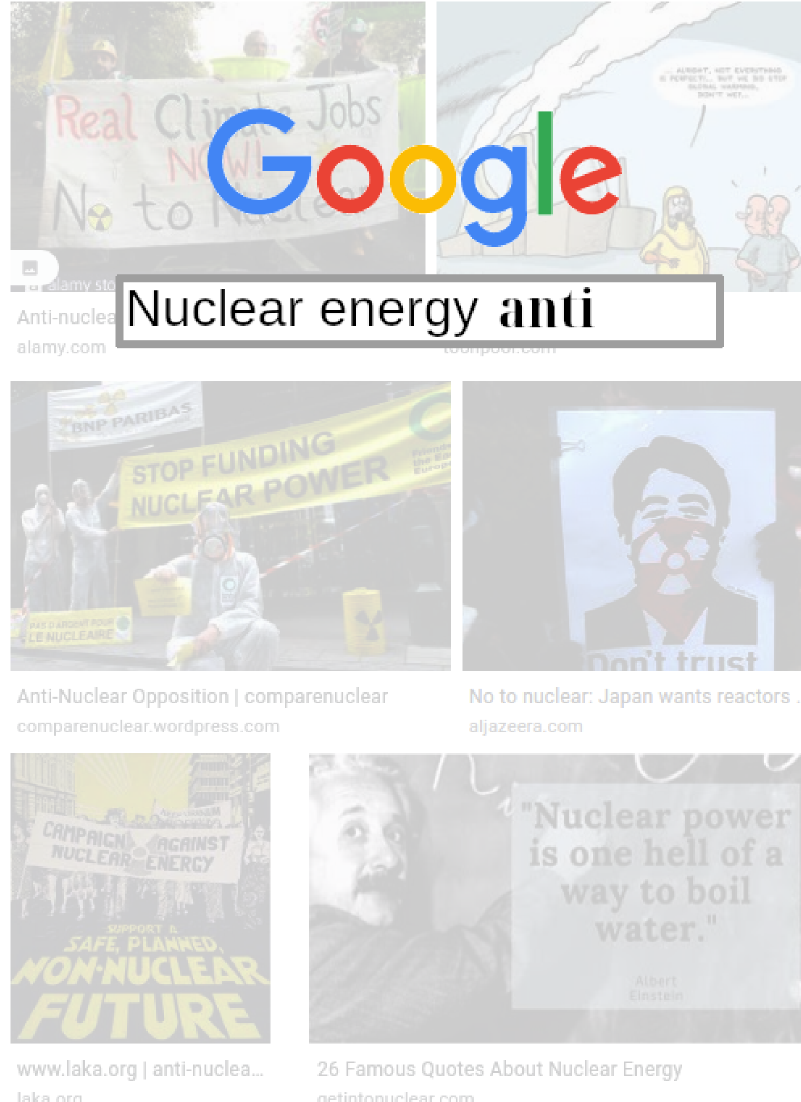
www.laka.org | anti-nuclea...
laka.org



26 Famous Quotes About Nuclear Energy
getintonuclear.com

Nuclear energy

- [x] cartoons
- [x] covers
- [x] photos
- [x] posters
- [x] memes
- [x] protesters
- [x] quotes
- [x] statistics

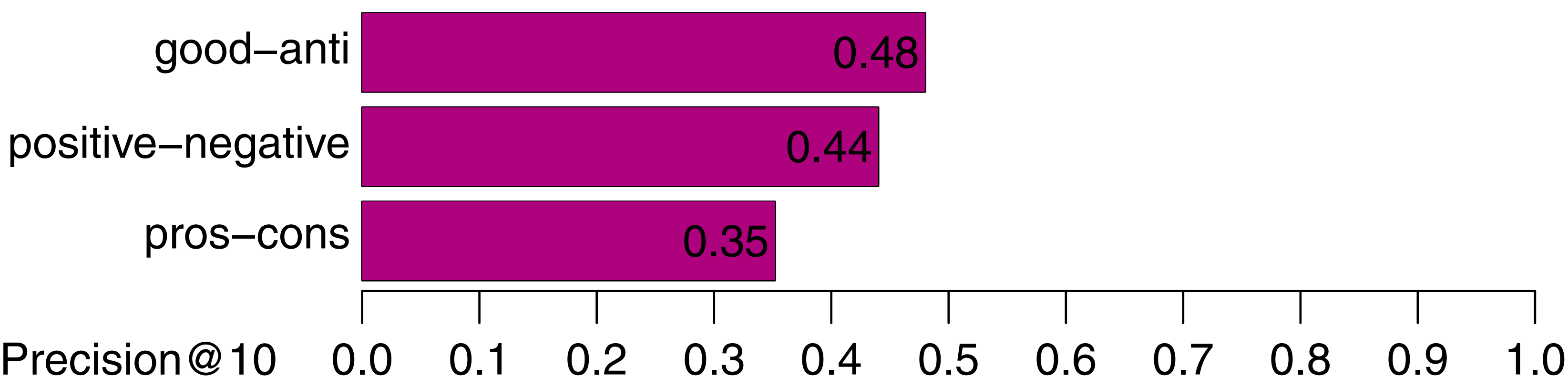


@KieselJohannes

Nuclear energy

☑ cartoons ☑ covers ☑ photos ☑ posters
☑ memes ☑ protesters ☑ quotes ☑ statistics

Evaluation on 20 Touché topics.



@KieselJohannes

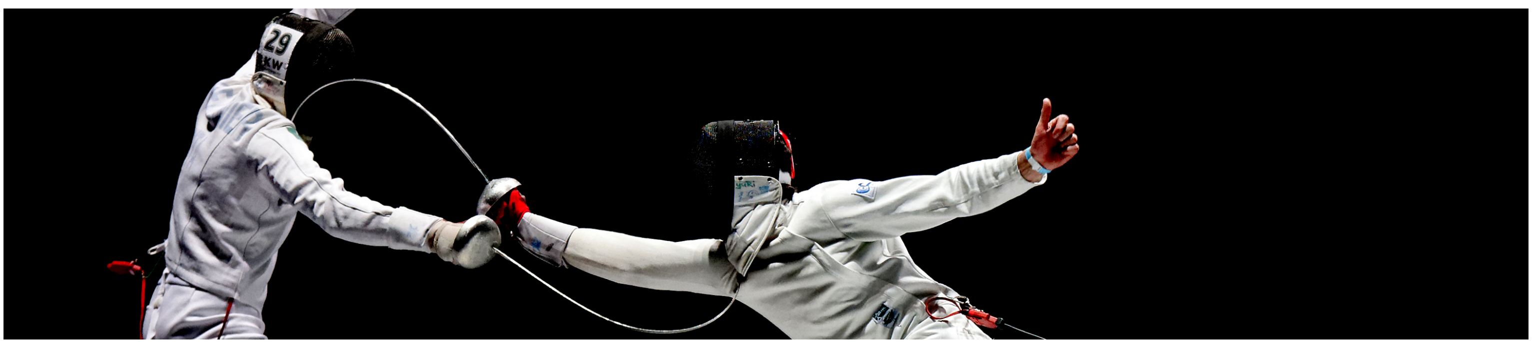
TOUCHÉ 2022 @ CLEF

Shared Tasks

- Task 1: Argument Retrieval for Controversial Questions.
- Task 2: Argument Retrieval for Comparative Questions.
- Task 3: **Image Retrieval for Arguments**.

Important Dates

- May 13, 2022: Approaches submission deadline.
- June 9, 2022: Participant paper submission.



More info now: touche.webis.de

Automatic summarization of German court decisions

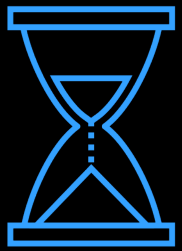
About the Project

- Topic of my masters thesis in cooperation with DATEV
- Goal: Summarize court decisions
  Foundation work, which will be continued in further projects

Why summarize court decisions?

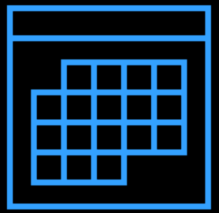
- Expand and interpret current laws
- Form the basis for forthcoming decisions
- Assessing the opinion of a judge or a senate

Why not summarize court decisions?



Very time consuming

Special knowledge required

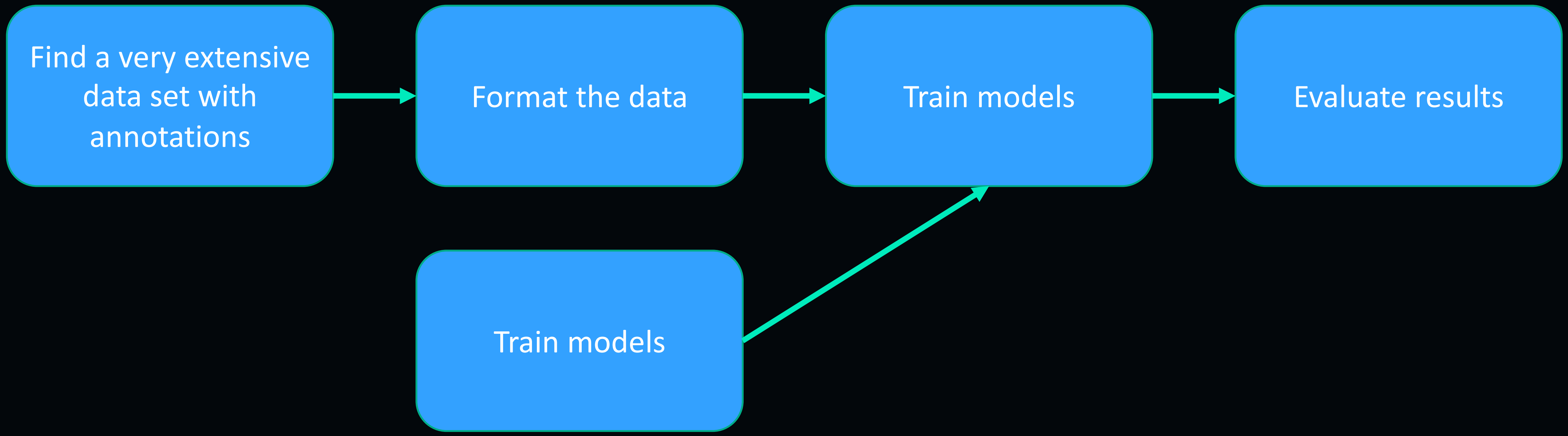


Ephemeral (relatively fast outdated)

Example

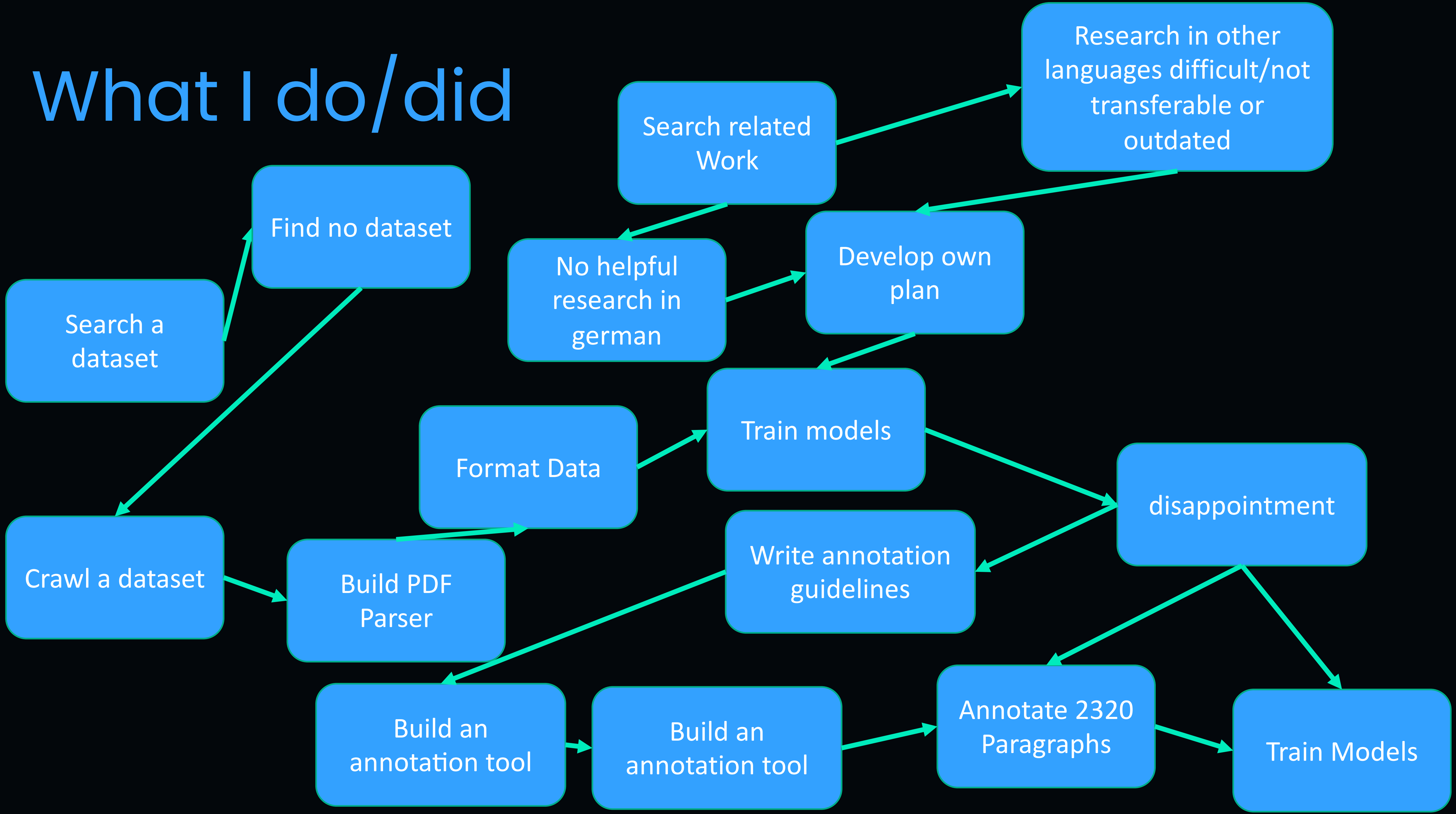
WHG § 3 Abs. 1 (§ 3 WHG - Einzelnorm, n.d.)	Sächsisches Wassergesetz: Kommentar für die Praxis (Dallhammer et al., 2019)
Für dieses Gesetz gelten folgende Begriffsbestimmungen: **Oberirdische Gewässer das ständig oder zeitweilig in Betten fließende oder stehende oder aus Quellen wild abfließende Wasser**	Unter einem Gewässerbett ist eine äußerlich erkennbare natürliche oder künstliche Begrenzung des Wassers in einer Eintiefung an der Erdoberfläche zu verstehen (vgl. BVerwG, Urt. v. 31.10.1975, BVerwGE Bd. 49 S. 293, 298; Beschl. v. 17.2.1969, Buchholz 445.4 § 1 WHG Nr. 3, m.w.N.). […] Von einem derartigen Bett kann u.a. dann nicht mehr gesprochen werden, wenn ein Graben vollständig verrohrt wird.

What I do/did



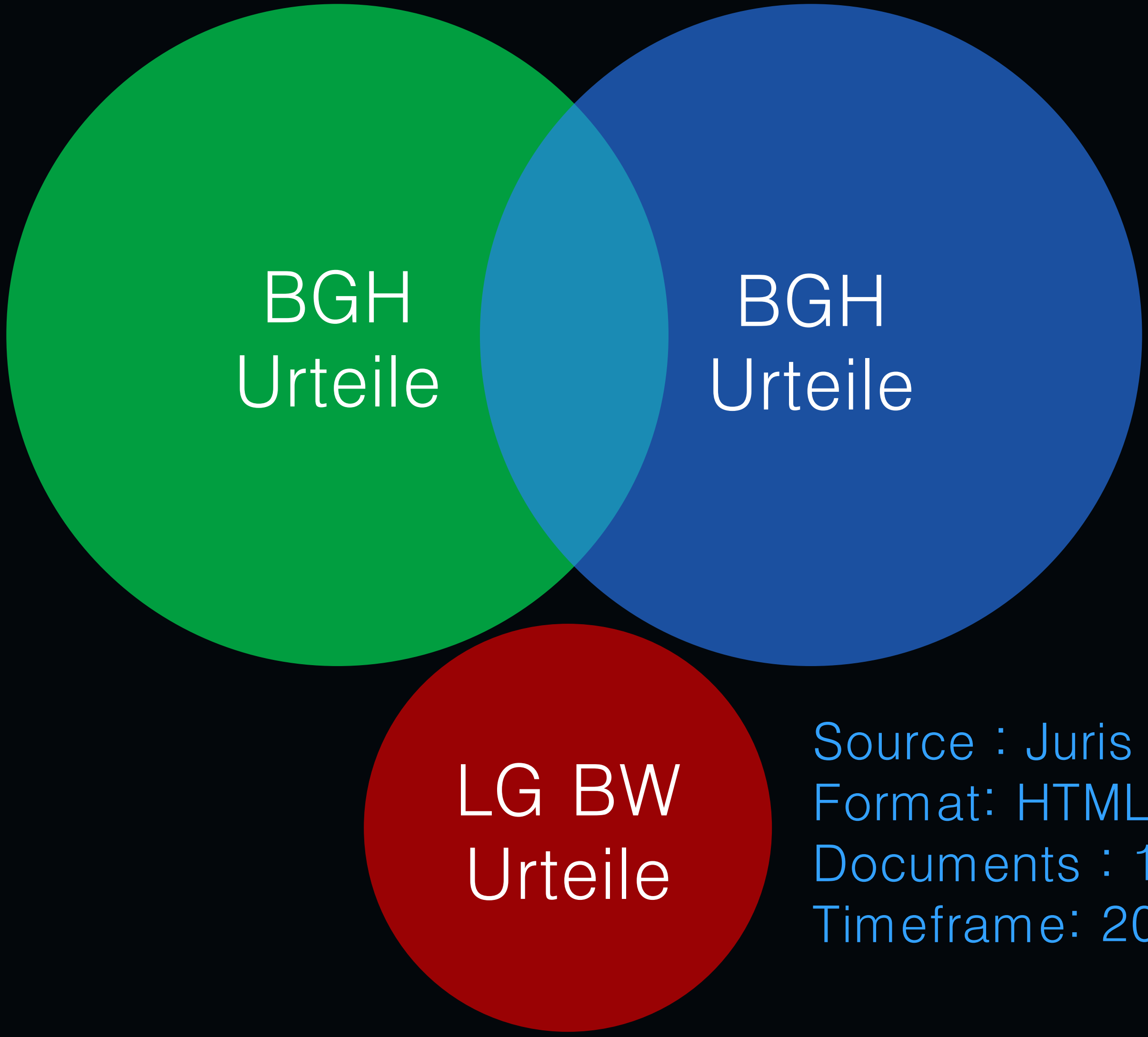
5 simple steps

What I do/did


Search a dataset
Find no dataset
Crawl a dataset
Build PDF Parser
Format Data
Search related Work
Research in other languages difficult/not transferable or outdated
No helpful research in german
Develop own plan
Train models
disappointment
Write annotation guidelines
Build an annotation tool
Build an annotation tool
Annotate 2320 Paragraphs
Train Models

Dataset

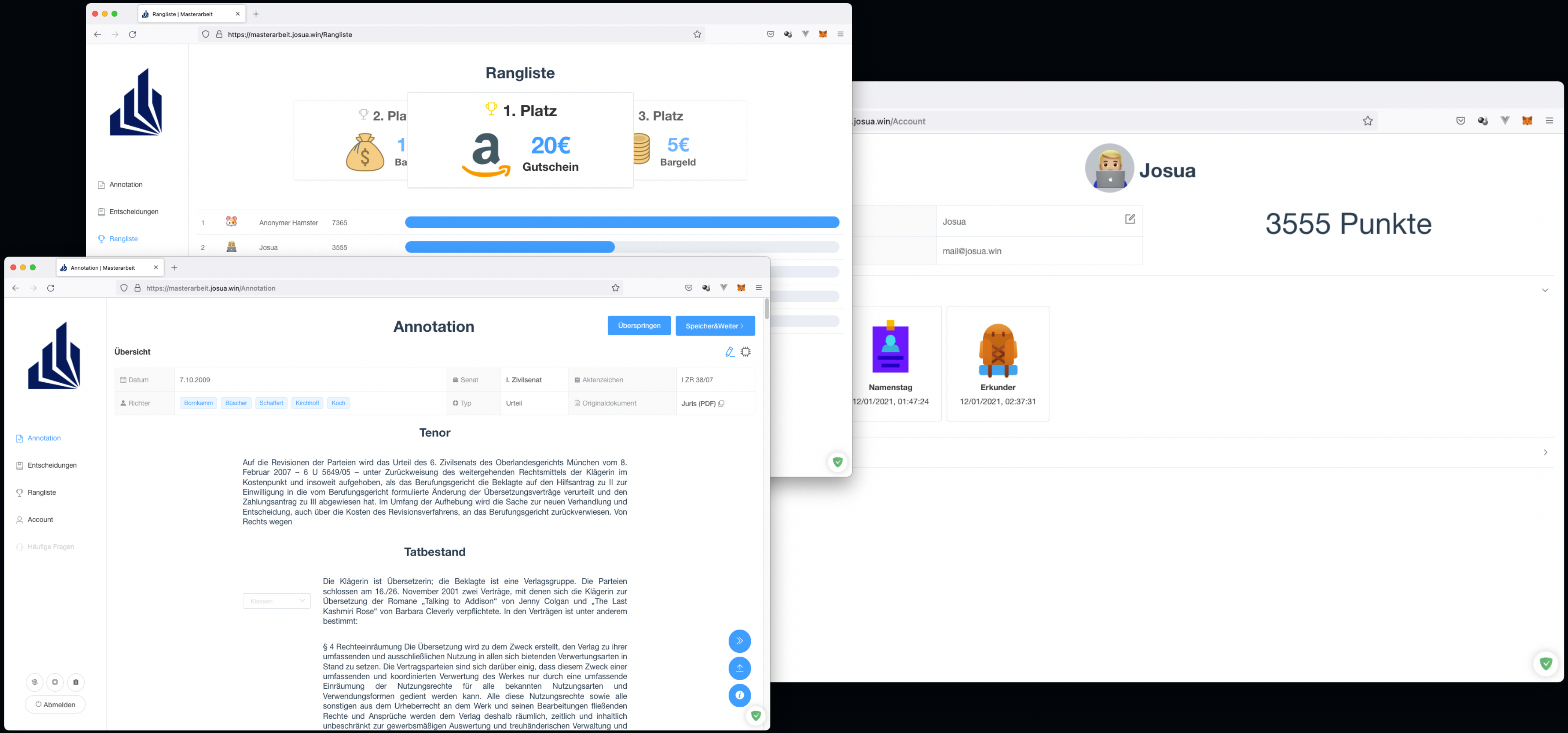
Source : DATEV
Format: XML
Documents : 53.730
Timeframe : 1970–today



Source : Juris
Format: PDF
Documents : 56.011
Timeframe: 2000–today

Source : Juris
Format: HTML
Documents : 18.758
Timeframe: 2006–today

Annotation Tool



GPT-3 Results

- Abstractive summarization
- Fine-Tuned on ~1000 documents
- Much better than all other GPTs (GPT-J, GPT-neo)
- Manual review by several lawyers confirmed the quality

Original	Der Revisionsbeklagte kann die Aufnahme eines im Revisionsverfahren durch Eröffnung des Insolvenzverfahrens unterbrochenen Prozesses…
Generted	Die Aufnahme eines durch die Eröffnung des Insolvenzverfahrens unterbrochenen Verfahrens unterliegt als Prozeßhandlung dem Anwaltszwang (§ 78 Abs. 1 ZPO); der nach § 250 ZPO einzureichende Schriftsatz muß von einem Rechtsanwalt unterschrieben sein. Vor den Gerichten des höheren Rechtszuges kann eine dem Anwaltszwang…

More examples: https://masterarbeit.josua.win/Ergebnisse

In Dialog with the Virtual Museum

Mariam Korashi (presenting)
Mervin D’Souza

Under Supervision
Johannes Kiesel
Marcel Gohsen

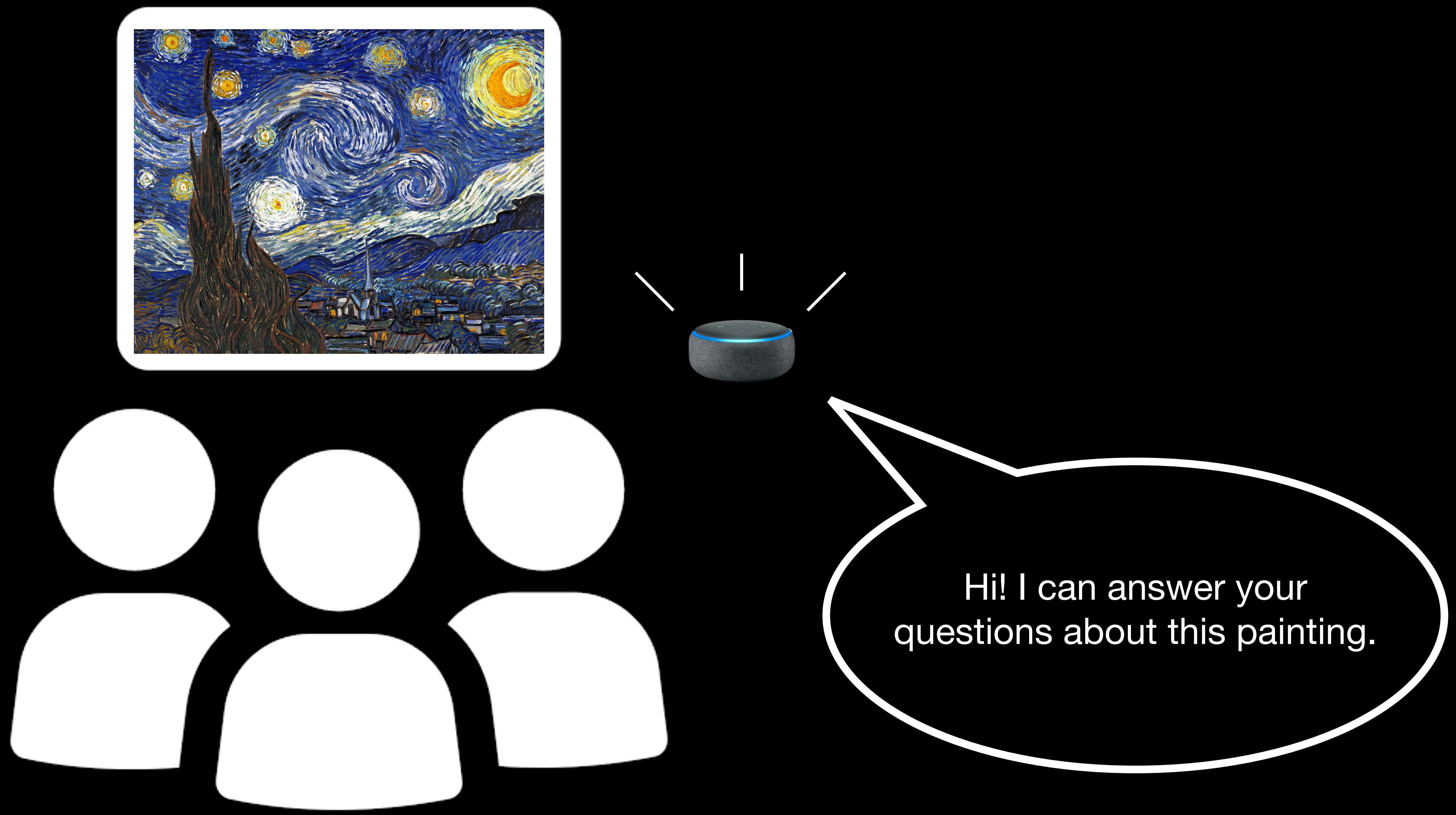
Bauhaus Universität Weimar

March 11, 2022 - Flash Talks 2022

Goal

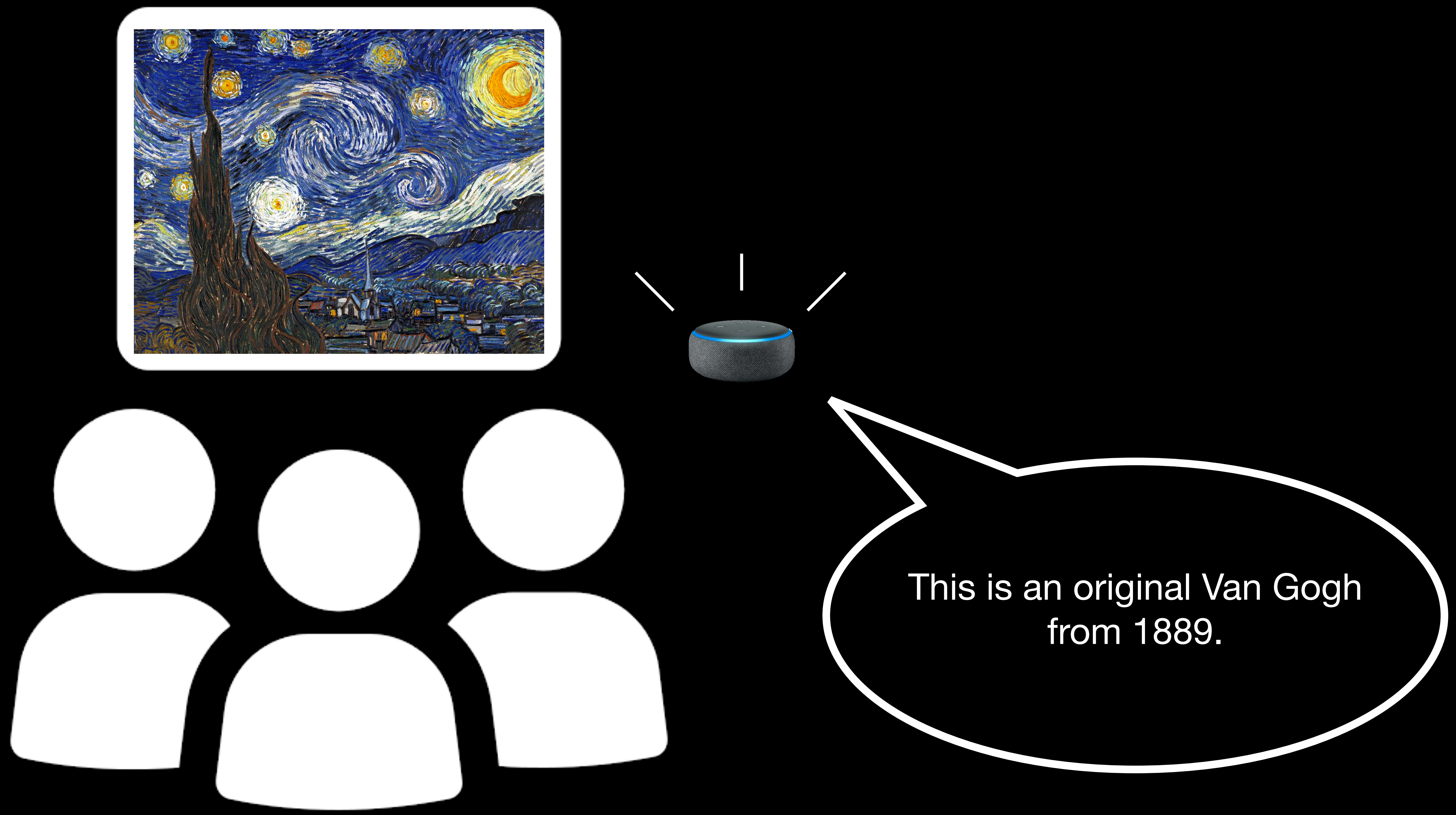
Training Alexa to answer questions in a (virtual) museum-context

Training Alexa to answer questions in a (virtual) museum-context



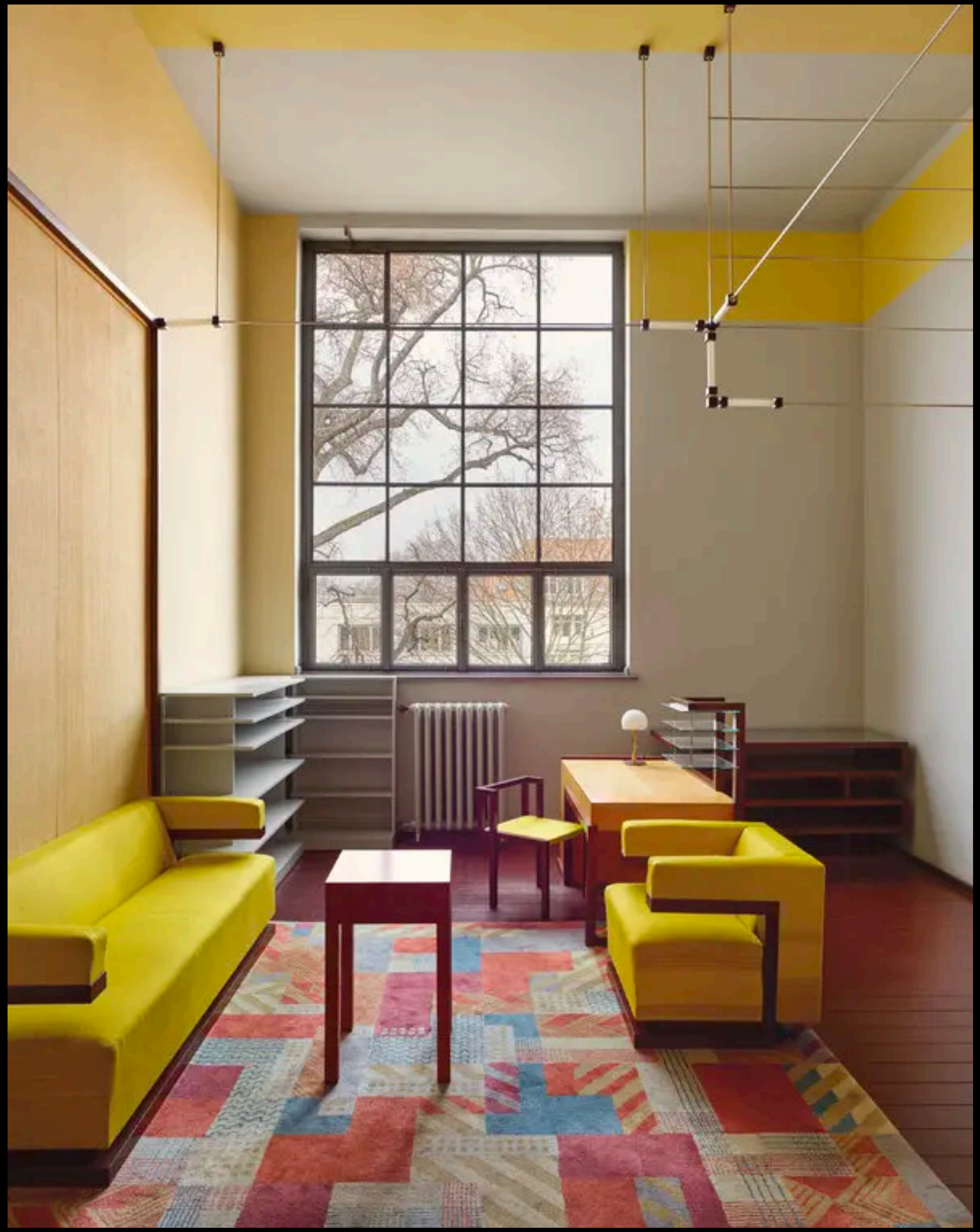
Training Alexa to answer questions in a (virtual) museum-context

Training Alexa to answer questions in a (virtual) museum-context



Goal

Training Alexa to answer questions in a ~~museum context~~ Virtual Museum


Gropius Room, Bauhaus Universität Weimar

Demo: Training Alexa to answer questions in a Virtual Museum

In

Goal

Training Alexa to answer questions in a museum-context

That’s simple.

Goal

Training Alexa to answer questions in a museum-context

That’s simple.

What is the difficult part?

Further Challenges

1. **How to keep the conversation going** (without overwhelming Alexa)
    - How to make users talk more/ask more questions

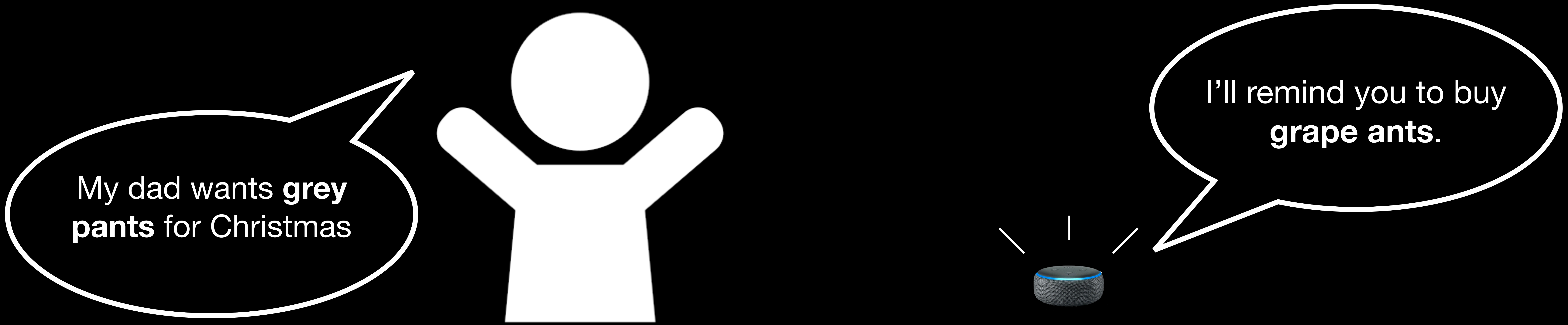
Further Challenges

1. **How to keep the conversation going** (without overwhelming Alexa)

Further Challenges

2. How users repair conversations

- Repeating? Re-explaining? Walk away?




THANK YOU

References

1. Icons from: thenounproject.com
2. Gropius Zimmer: https://fabricefouillet.com/galerie/bauhaus-centenary-t-magazine/#1

Identifying Exaggerated Language

Alonso Palomino

Identifying Exaggerated Language

What is exaggerated language?

- Exaggerated language
  - Expressions that improve or diminish the perception of an object or situation more or less than it actually is.
    - **Overstatements:** The act of stating something more profound to make the point more serious or important or beautiful. Writers use overstatement as a literary technique for the sake of humor and for emphasizing a certain point [1].
      - *In teaching his philosophy to others, Socrates never gave answers but only questioned people.*
    - **Hyperboles:** Boldly overstated or exaggerated claim or statement that adds emphasis without the intention of being literally true [2].
      - *He feels buried under a mountain of work.*

Identifying Exaggerated Language
Why studying it?

- **In argumentation:**
    - Hyperbolic statements can be used as rhetorical devices to praise someone or approve something. Also, to complain, attack, and criticize someone or something.
    - Persuasive purposes: hyperbolic statements function as an emphasizing (or bias) device and can therefore be used to make information that is to the arguer's advantage seem more salient to the audience.
    - Evaluative meaning: convey and arise emotions
- **It is a proxy for evaluating the credibility of a statement**
    - Hyperbole is a rhetorical trope through which statements are made that are obviously exaggerated and thus untrue or unwarranted [3].
    - Social media users use exaggerated language to make others see a phenomenon from their perspective, emphasizing one element, aspect, or account of a topic. For instance, TV information, debates, and verbal productions can contain overstatements that puff up facts [4].
- **Biased language**
    - Exaggeration could be used to skew facts deliberately.

Exaggerated Language

How to identify exaggerated language?

- Previous approaches to identify hyperbolic expressions rely on different approaches that harness:
  - Numerical expressions
  - Expressions of spatial extent
  - Intensifiers
  - Extreme adjectives and adverbs
  - Comparatives and superlatives
  - Polarity features
  - Imageability features

Exaggerated Language

Questions

- How do people exaggerate in argumentation?
  - What are the linguistic structures that people use to exaggerate?
- How to detect exaggerated expressions without relying on external linguistic resources?
- Is it possible to identify exaggerated utterances by utilizing syntactic, semantic, morphologic, and pragmatical features?
  - What linguistic information about exaggerated language is captured in neural networks?
  - How can attention models be utilized to investigate what specific aspects make an expression exaggerated?

References

1. https://literarydevices.net/overstatement/
2. https://literarydevices.net/hyperbole/
3. Snoeck Henkemans, A. F. (2013, May). The use of hyperbole in the argumentation stage. In International Conference of the Ontario Society for the Study of Argumentation (OSSA) (Vol. 22, p. 26).
4. Troiano, E., Strapparava, C., Özbal, G., & Tekiroğlu, S. S. (2018). A computational exploration of exaggeration. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (pp. 3296-3304).

≥ ir_axioms

Intuitive Axiomatic Retrieval Experimentation.

A. Bondarenko M. Fröbe J. H. Reimer B. Stein M. Völske M. Hagen

Axioms

What they promise 😉

- Axioms are (pairwise) **preferences**

≈ "Rules" what documents to rank first

- **Easy** to understand

Use-cases

- Explain ranker decisions
- Evaluate ranking errors
- Re-rank

Working with axioms… 😐

in the past

PROX1 [17] "Prefer the document with shorter total distance between query term pairs."

Given $|Q| > 1, \forall_{q \in Q} q \in D_1 \wedge q \in D_2, \quad M(D,q) = \{i : t_i \in D \wedge t_i = q\}$

$\delta(D, q_1, q_2) = \frac{1}{M(D,q_1) \cdot M(D,q_2)} \sum_{(i,j) \in M(D,q_1) \times M(D,q_2)} |i - j|$

$$\sum_{(q_i,q_j) \in Q \times Q} \delta(D_1, q_i, q_j) < \sum_{(q_i,q_j)} \delta(D_2, q_i, q_j) \quad \Rightarrow \quad D_1 >_{\text{PROX1}} D_2$$

PROX2 [17] "Prefer documents where query terms occur earlier."

Given $|Q| > 1, \forall_{q \in Q} q \in D_1 \wedge q \in D_2, \quad \text{first}(q, D) = \min\{i : t_i \in D \wedge t_i = q\}$

$$\sum_{q \in Q} \text{first}(q, D_1) < \sum_{q \in Q} \text{first}(q, D_2) \quad \Rightarrow \quad D_1 >_{\text{PROX2}} D_2$$

PROX3 [17] "Prefer documents where the query occurs earlier as a phrase."

Given $Q = \{q_1, \dots, q_l\}, \forall_{q \in Q} q \in D_1 \wedge q \in D_2,$

$\tau(Q, D) = \min\{i : t_i \in D \wedge t_i = q_1, \dots, t_{i+l} = q_l\} \cup \{\infty\}$

$$\tau(D_1, Q) < \tau(D_2, Q) \quad \Rightarrow \quad D_1 >_{\text{PROX3}} D_2$$

PROX4 [17] "Prefer documents that cover all query terms in a shorter sub-string."

Given $|Q| > 1, \forall_{q \in Q} q \in D_1 \wedge q \in D_2,$

$\omega(D, Q) = \min\{j - i : i < j \wedge t_i \in D \wedge t_j \in D \wedge \forall_{q \in Q} q \in D_{|i \dots j}\}$

$$\omega(D_1, Q) < \omega(D_2, Q) \quad \Rightarrow \quad D_1 >_{\text{PROX4}} D_2$$

PROX5 [17] "Prefer documents where query terms are closer together on average."

Given $|Q| > 1, \forall_{q \in Q} q \in D_1 \wedge q \in D_2, \quad M(D, Q) = \{i : t_i \in D \wedge t_i \in Q\}$

$s(D, Q, i) = \min\{k - j : j \leq i \wedge k \geq i \wedge \forall_{q \in Q} q \in D_{|j \dots k}\}$

$$\frac{\sum_{i \in M(D_1,Q)} s(D_1,Q,i)}{|M(D_1,Q)|} < \frac{\sum_{i \in M(D_2,Q)} s(D_2,Q,i)}{|M(D_2,Q)|} \quad \Rightarrow \quad D_1 >_{\text{PROX5}} D_2$$

Problems

- Not always "easy to understand"™
- Many implementation caveats (uff!)
- Hard to maintain etc.…

`ir_axioms` to the rescue!

`webis-de/ir_axioms` `pip install ir_axioms`

- Reference implementations for **25 common axioms**
- **Define and combine** axioms declaratively
- Tightly integrates with **PyTerrier & Pyserini**

Experiments

```
experiment = AxiomaticExperiment(
  [bm25, monot5, ...],
  dataset.get_topics(),
  dataset.get_qrels(),
  index,
  axioms=[ArgUC(), QTArg(), QTPArg(), ...]
)

experiment.preferences
experiment.preference_distribution
experiment.preference_consistency
experiment.inconsistent_pairs
```

Re-ranking

```
bm25 = BatchRetrieve(index, "BM25")

axiom = (ArgUC() & QTArg() & QTPArg()) | ORIG()

# Re-rank top-20 documents with KwikSort.
kwiksort = bm25 % 20 >> \
  KwikSortReranker(axiom, index)

pipeline = kwiksort ^ bm25
```

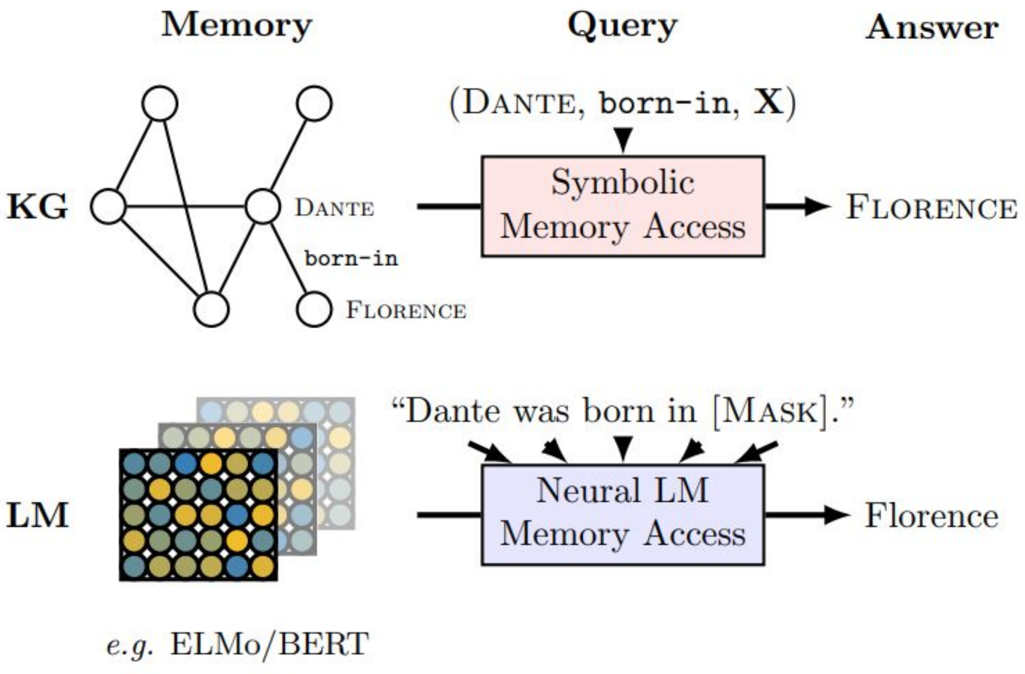
Thank you, stay tuned for more!

Train BERT to Know. Better.

Learning World Knowledge from Knowledge Graphs

Erik Reuter, 11.03.2022

BERT Knows: Language Models as Knowledge Bases?



Petroni et al.
Language Models as Knowledge Bases?, published in CoRR, 2019

BERT Knows… Not That Well

T-Rex Relation Type	KG String-Matching (Precision)	BERT-Base (Precision)	BERT-Large (Precision)
1-1 Relations	10.0	68.0	**74.5**
N-1 Relations	33.8	32.4	**34.2**
N-M Relations	**36.7**	24.7	24.3
Total	**33.8**	31.1	32.3

Petroni et al.
Language Models as Knowledge Bases?, published in CoRR, 2019

T-Rex Dataset

- Passages with annotated knowledge-base triples (head, relation, tail)

“Spongiforma squarepantsii is a species of fungus in the family Boletaceae.”

Elsahar et al.
T-REx: A Large Scale Alignment of Natural Language with Knowledge Base Triples, accepted at LRAC, 2018

T-Rex Dataset

- Passages with annotated knowledge-base triples (head, relation, tail)

"Spongiforma squarepantsii is a species of fungus in the family Boletaceae."

Elsahar et al.
T-REx: A Large Scale Alignment of Natural Language with Knowledge Base Triples, accepted at LRAC, 2018

T-Rex Dataset

- Passages with annotated knowledge-base triples (head, relation, tail)

“Spongiforma squarepantsii is a species of fungus in the family Boletaceae.”

Elsahar et al.
T-REx: A Large Scale Alignment of Natural Language with Knowledge Base Triples, accepted at LRAC, 2018

T-Rex Dataset

- Passages with annotated knowledge-base triples (head, relation, tail)

"Spongiforma squarepantsii is a species of fungus in the family Boletaceae."

Elsahar et al.
T-REx: A Large Scale Alignment of Natural Language with Knowledge Base Triples, accepted at LRAC, 2018

T-Rex Dataset

- Passages with annotated knowledge-base triples (head, relation, tail)

“Spongiforma squarepantsii is a species of fungus in the family Boletaceae.”

-> (Spongiforma squarepantsii, is a, species of fungus)

Elsahar et al.
T-REx: A Large Scale Alignment of Natural Language with Knowledge Base Triples, accepted at LRAC, 2018

BERT Knows… With KG-Embeddings

- Use annotated triples for knowledge embedding:

"Spongiforma squarepantsii is a species of fungus in the family Boletaceae."

-> (Spongiforma squarepantsii, is a, species of fungus)

-> In the embedding space: e(S. squarepantsii) + e(is a) = e(species of fungus)

BERT Knows… With KG-Embeddings

- Use annotated triples for knowledge embedding:

"Spongiforma squarepantsii is a species of fungus in the family Boletaceae."

-> (Spongiforma squarepantsii, is a, species of fungus)

-> In the embedding space: e(S. squarepantsii) + e(is a) = e(species of fungus)

- Naive loss function: ||e(head) + e(relation) - e(tail)||

Conclusion and Next Steps

- BERT knows, but not well enough
- Using knowledge graphs to transfer knowledge into BERT
    - T-Rex combines passages with triples
    - Train via masked language modeling and knowledge graph loss simultaneously

- Next steps:
    - Test different knowledge graph loss functions
    - Evaluate different loss functions

Thank you!

Sources

Sponge Wikipedia Page: https://en.wikipedia.org/wiki/Spongiforma_squarepantsii , accessed 18:01 CET

Elsahar et al., T-REx: A Large Scale Alignment of Natural Language with Knowledge Base Triples, accepted at LRAC, 2018, http://www.lrec-conf.org/proceedings/lrec2018/summaries/632.html

Petroni et al., Language Models as Knowledge Bases?, published in CoRR, 2019, http://arxiv.org/abs/1909.01066

SpongeBob:
https://vignette.wikia.nocookie.net/spongebob/images/5/5c/Spongebob-squarepants.png/revision/latest?cb=20190424195119&path-prefix=de

Toad:

https://www.netclipart.com/pp/m/109-1097875_super-mario-odyssey-toad-red-toad-super-mario.png

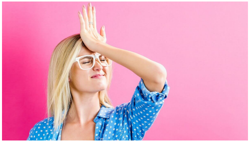
Simulation of False Memories in Known-Item Searches

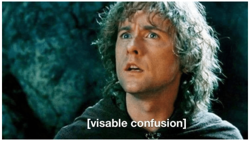
Matthias Hagen Maik Fröbe Eric Schmidt

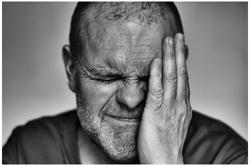
Martin-Luther-Universität Halle-Wittenberg

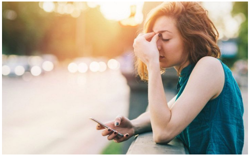

ME TRYING TO REMEMBER
WHAT I FORGOT
ABOUT THAT THING I FORGOT
imgflip.com

This is not only a problem of Professor X!









What we want to achieve

Kiade

asked in Entertainment & Music > Movies

Do You Know The Name of this 90's Movie?

It's a pretty old movie about this man, I'm pretty sure he's a criminal, and this little boy, I think he kidnaps him, and he ends up taking care of him all the way until the end of the film. He dies by a tree as the police are coming.
I think the main guy is Bruce Willis.

Answer Save

What we want to achieve

```
90s crime movie young boy hostage Bruce Willis
```

What we want to achieve

`90s crime movie young boy hostage Bruce Willis`

Hostage (2005)

What we want to achieve

false memory:

`90s crime movie young boy hostage Bruce Willis`

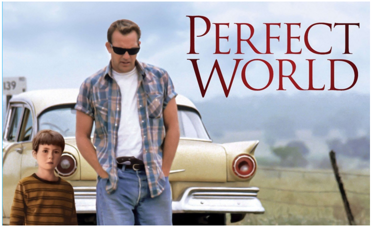
Hostage (2005)

What we want to achieve

false memory:

`90s crime movie young boy hostage Bruce Willis`

Hostage (2005)



A Perfect World (1993)

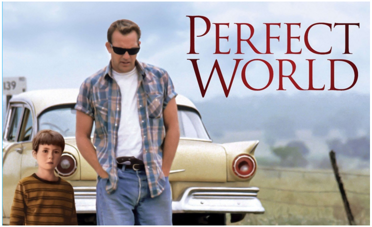
`90s crime movie young boy hostage Kevin Costner`

What we want to achieve

false memory:

`90s crime movie young boy hostage Bruce Willis`

Hostage (2005)



A Perfect World (1993)

`90s crime movie young boy hostage Kevin Costner`

corrected search query

False Memories from Yahoo!-Answers

extracted from a Yahoo!-Answers dataset

- 183.824.847 question and answer pairs
- 27 TB

extracted with 97 rules. Example:

- answer contains: `It's not ..., but ...`

Category	Proportion
false memories (overall)	8,7 %

False Memories from Yahoo!-Answers

extracted from a Yahoo!-Answers dataset

- 183.824.847 question and answer pairs
- 27 TB

extracted with 97 rules. Example:

- answer contains: `It's not ..., but ...`

Category	**Proportion**
false memories (overall)	8,7 %
false memories (optimized rules, books)	**24,0 %**

False Memories from Yahoo!-Answers

extracted from a Yahoo!-Answers dataset

- 183.824.847 question and answer pairs
- 27 TB

extracted with 97 rules. Example:

- answer contains: `It's not ..., but ...`

Category	**Proportion**
false memories (overall)	8,7 %
false memories (optimized rules, books)	**24,0 %**
false memories (optimized rules, movies)	**35,0 %**

Simulate False Memories

Kevin Costner (Q11930)

American actor and filmmaker
Kevin Michael Costner

Statements

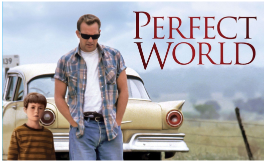
instance of	human 3 references	edit + add value
sex or gender	male 5 references	edit + add value
country of citizenship	United States of America 0 references	edit + add reference + add value
date of birth	18 January 1955 8 references	edit + add value

Simulate False Memories

`90s crime movie young boy hostage ...`

`Bruce Willis` `Kevin Costner`

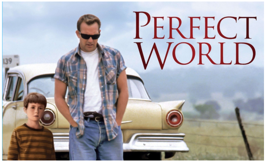
Hostage (2005)



A Perfect World (1993)

Simulate False Memories

`90s crime movie young boy hostage ...`

`Bruce Willis`	`Kevin Costner`	`Michael Keaton`
		
Hostage (2005)	A Perfect World (1993)	Desperate Measures (1998)

Impact of False Memories on Search Results

Search query	∅ Position
real false memories from Yahoo!-Answers	20
correct Yahoo!-Answers questions	**4**
simulated false memories	19
simulated correct keyquery	**3**

based on 18 Yahoo!-Answers topics

Achieved so far

- Extract and analyse false memories from Yahoo!-Answers
- Simulate false memories
- Evaluate the impact of false memories on search results

Next Steps

- Measure similarities of two entities: Embedding Vectors
- Label more false memories using crowdsourcing
- Detect and correct false memories in search queries

Achieved so far

- Extract and analyse false memories from Yahoo!-Answers
- Simulate false memories
- Evaluate the impact of false memories on search results

Next Steps

- Measure similarities of two entities: Embedding Vectors
- Label more false memories using crowdsourcing
- Detect and correct false memories in search queries

Achieved so far

- Extract and analyse false memories from Yahoo!-Answers
- Simulate false memories
- Evaluate the impact of false memories on search results

Next Steps

- Measure similarities of two entities: Embedding Vectors
- Label more false memories using crowdsourcing
- Detect and correct false memories in search queries

Academic Self-Presentation on the Web

Towards insights in Bibliometrics from causal models

Master’s thesis, CS4DM - Sagar Nagaraj Simha

Supervisors - Dr. Magdalena Anna Wolska, Prof. Benno Stein,
Prof. Jakob Runge (DLR)

Who is a researcher?

Qualitative researcher

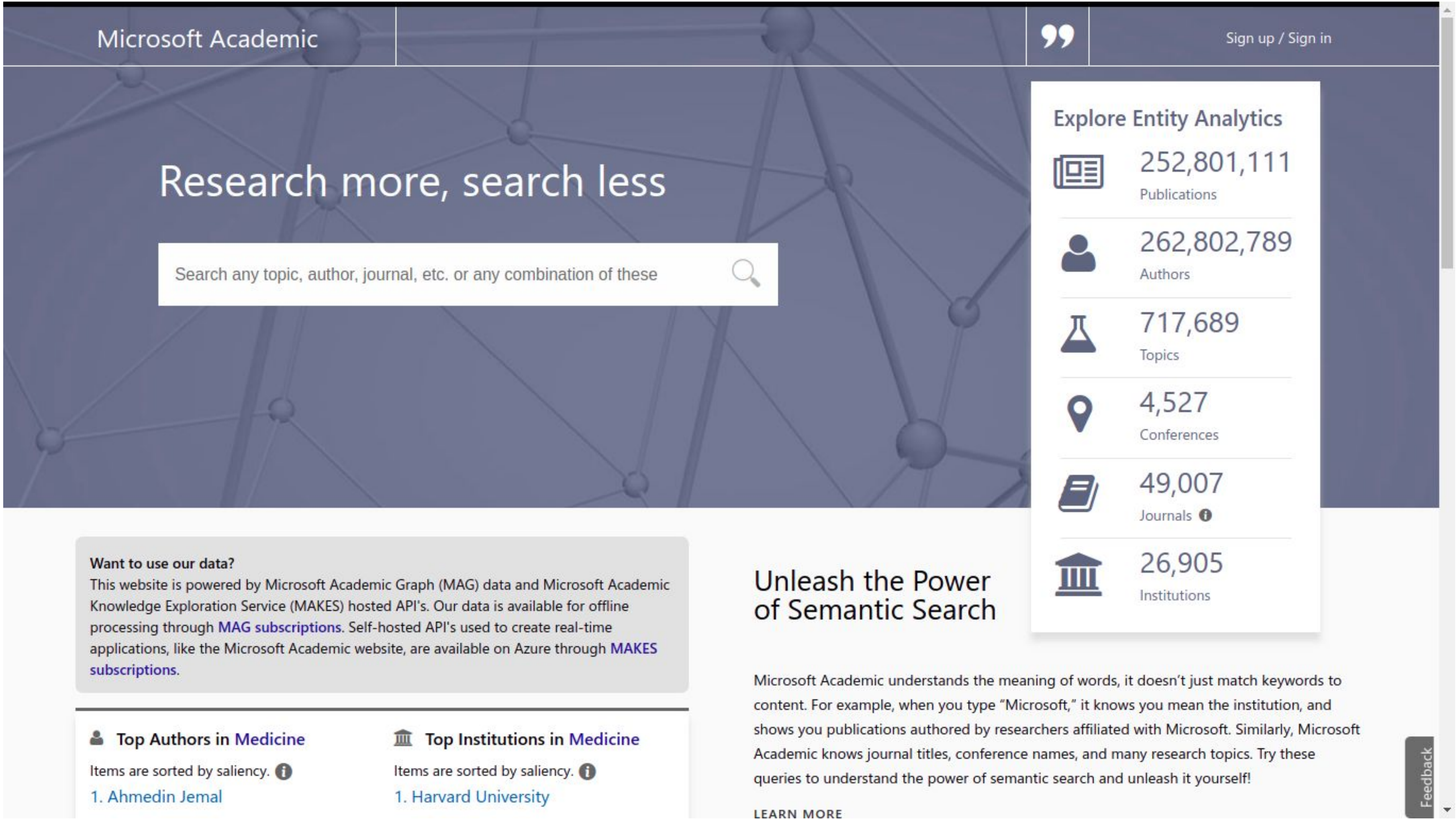
Empirical researcher

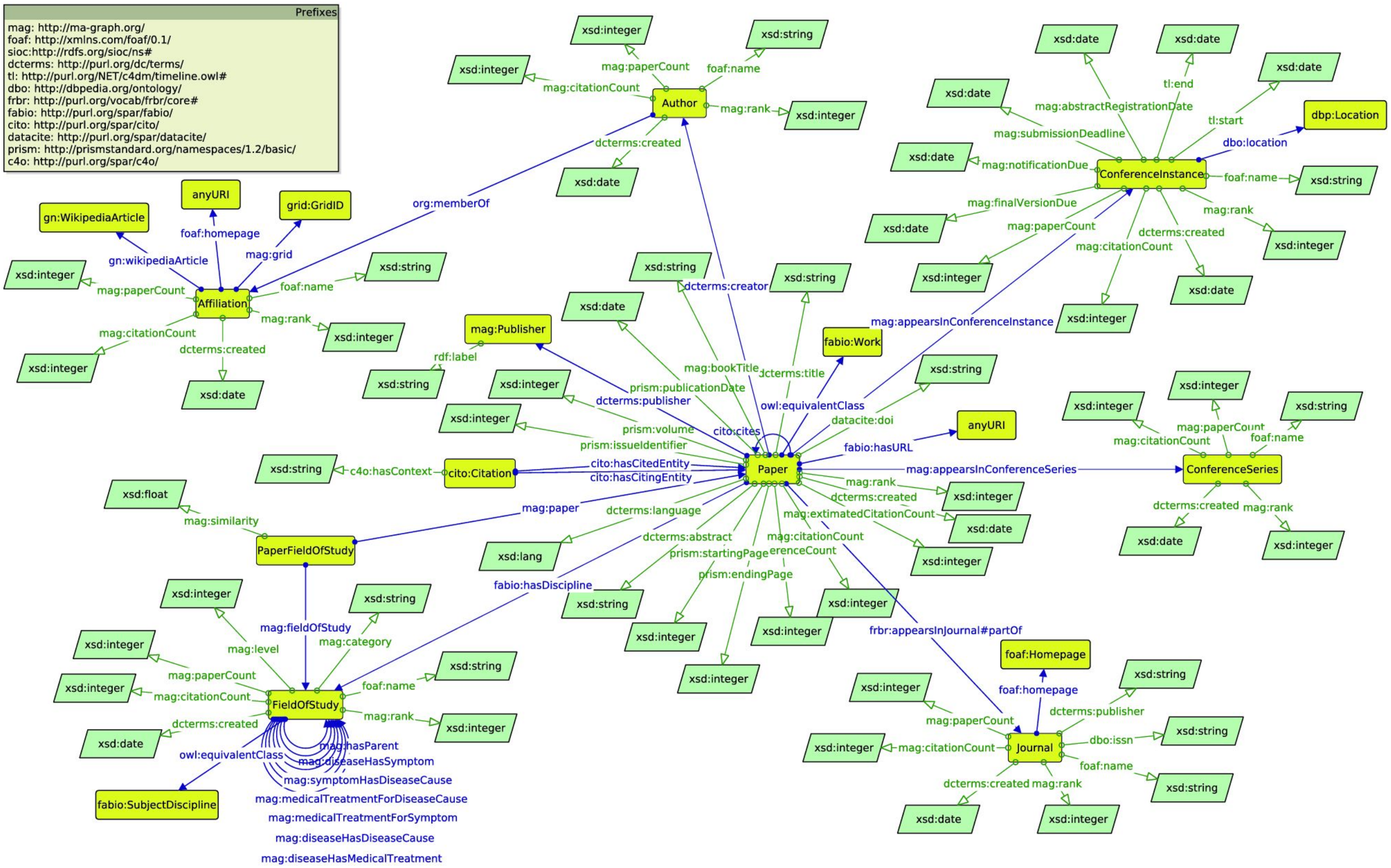
The Four Types of Scientists - National Geographic (opinion piece)

- The Data-Driven Nerd/Myers-Briggs ISTJ or ISTP
- The Theory-Driven Nerd/Myers-Briggs INTJ or INTP
- The Data-Driven Adventurer/Myers-Briggs ESTP
- The Theory-Driven Adventurer/Myers-Briggs ENTP

Can publications of a researcher tell us what kind they are?

Analysis using Bibliometric features

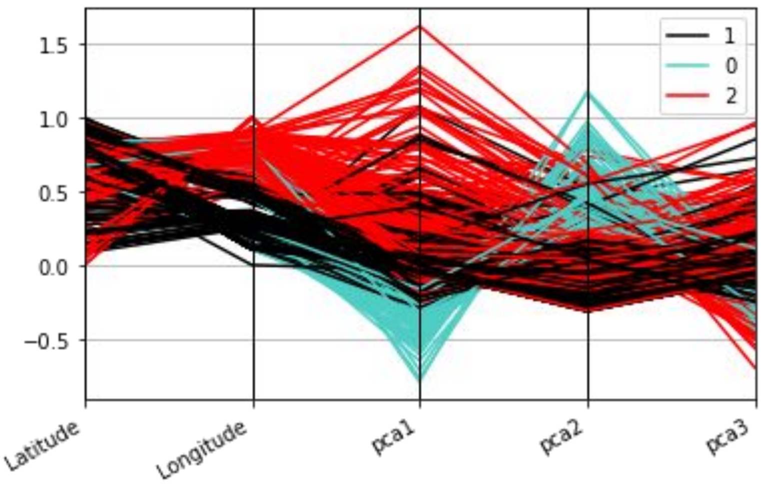

Microsoft Academic
Sign up / Sign in
Research more, search less
Search any topic, author, journal, etc. or any combination of these
Explore Entity Analytics
252,801,111
Publications
262,802,789
Authors
717,689
Topics
4,527
Conferences
49,007
Journals
26,905
Institutions
Want to use our data?
This website is powered by Microsoft Academic Graph (MAG) data and Microsoft Academic Knowledge Exploration Service (MAKES) hosted API's. Our data is available for offline processing through MAG subscriptions. Self-hosted API's used to create real-time applications, like the Microsoft Academic website, are available on Azure through MAKES subscriptions.
Top Authors in Medicine
Items are sorted by saliency.
1. Ahmedin Jemal
Top Institutions in Medicine
Items are sorted by saliency.
1. Harvard University
Unleash the Power of Semantic Search
Microsoft Academic understands the meaning of words, it doesn't just match keywords to content. For example, when you type "Microsoft," it knows you mean the institution, and shows you publications authored by researchers affiliated with Microsoft. Similarly, Microsoft Academic knows journal titles, conference names, and many research topics. Try these queries to understand the power of semantic search and unleash it yourself!
LEARN MORE
Feedback

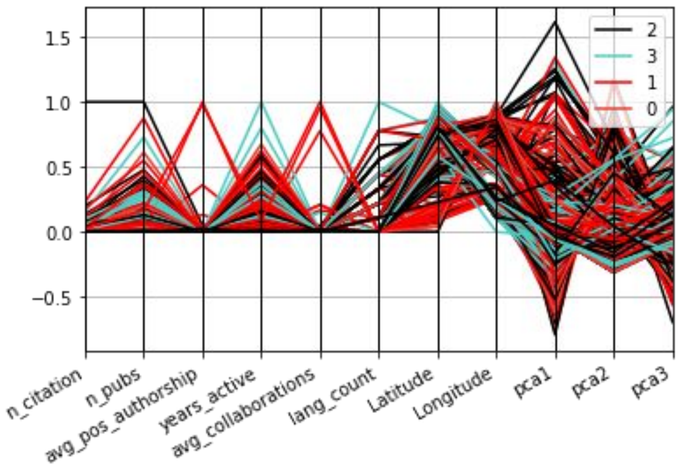


Prefixes
mag: http://ma-graph.org/
foaf: http://xmlns.com/foaf/0.1/
sioc:http://rdfs.org/sioc/ns#
dcterms: http://purl.org/dc/terms/
tl: http://purl.org/NET/c4dm/timeline.owl#
dbo: http://dbpedia.org/ontology/
frbr: http://purl.org/vocab/frbr/core#
fabio: http://purl.org/spar/fabio/
cito: http://purl.org/spar/cito/
datacite: http://purl.org/spar/datacite/
prism: http://prismstandard.org/namespaces/1.2/basic/
c4o: http://purl.org/spar/c4o/
Author
Affiliation
Paper
ConferenceInstance
ConferenceSeries
Journal
FieldOfStudy
PaperFieldOfStudy
cito:Citation
mag:Publisher
fabio:Work
fabio:SubjectDiscipline
foaf:Homepage
dbp:Location
gn:WikipediaArticle
grid:GridID
anyURI
org:memberOf
dcterms:creator
mag:appearsInConferenceInstance
mag:appearsInConferenceSeries
frbr:appearsInJournal#partOf
fabio:hasDiscipline
mag:paper
mag:fieldOfStudy
cito:hasCitedEntity
cito:hasCitingEntity
c4o:hasContext
cito:cites
owl:equivalentClass
fabio:hasURL
dcterms:publisher
prism:publicationDate
mag:bookTitle
dcterms:title
datacite:doi
prism:volume
prism:issueIdentifier
mag:rank
dcterms:created
mag:extimatedCitationCount
mag:citationCount
erenceCount
dcterms:language
dcterms:abstract
prism:startingPage
prism:endingPage
mag:paperCount
foaf:name
mag:abstractRegistrationDate
tl:end
tl:start
mag:submissionDeadline
mag:notificationDue
mag:finalVersionDue
dbo:location
dbo:issn
foaf:homepage
gn:wikipediaArticle
mag:grid
rdf:label
mag:similarity
mag:level
mag:category
mag:hasParent
mag:diseaseHasSymptom
mag:symptomHasDiseaseCause
mag:medicalTreatmentForDiseaseCause
mag:medicalTreatmentForSymptom
mag:diseaseHasDiseaseCause
mag:diseaseHasMedicalTreatment
xsd:integer
xsd:string
xsd:date
xsd:float
xsd:lang

Distilled features

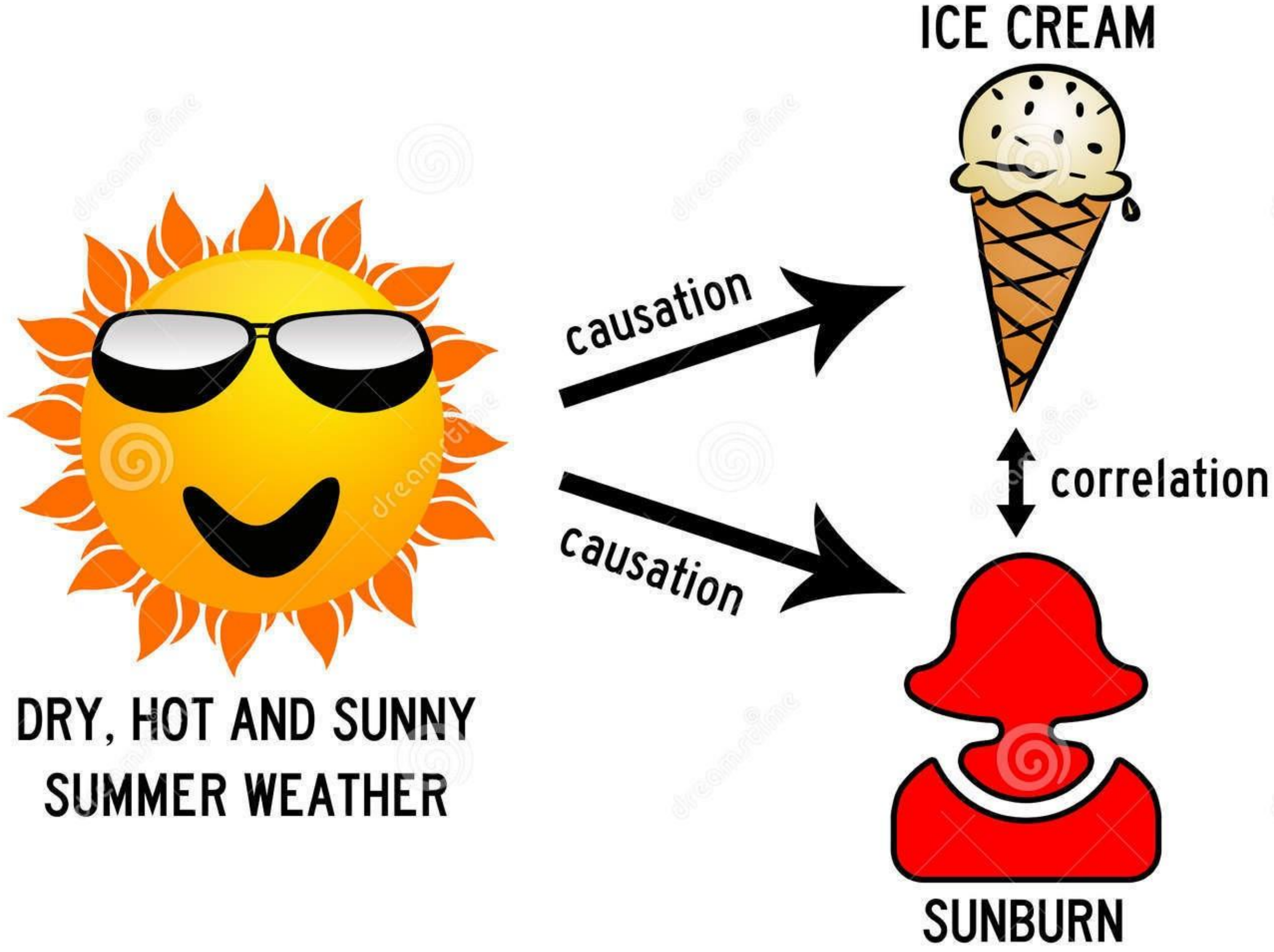
- Number of citations - 'n_citation'
- Number of publications - 'n_pubs',
- Affiliation name - 'affiliation_name' (hash-encoded)
- **Title (Principal components of cosine similarity based on TFIDF)**
- **Location - 'Latitude', 'Longitude'**
- Average authorship position - 'avg_pos_authorship'
- Years active in publication - 'years_active'
- Average collaborations over publications - 'avg_collaborations'
- Number of languages published in - 'lang_count'





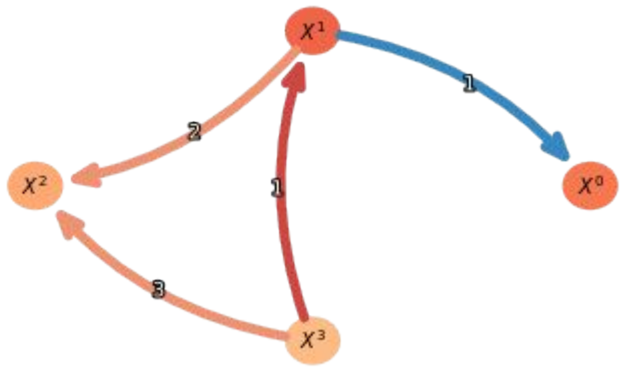
Each line represents one author

Can Causal models help identify causal features?


ICE CREAM
causation
causation
correlation
DRY, HOT AND SUNNY SUMMER WEATHER
SUNBURN

Causal inference and causal discovery

- Tigramite - https://github.com/jakobrunge/tigramite
- Data preprocessing - generates sets of (X, Y, Z)
- **Conditional independence testing**
- Causal discovery - Ex : PCMCI algorithm



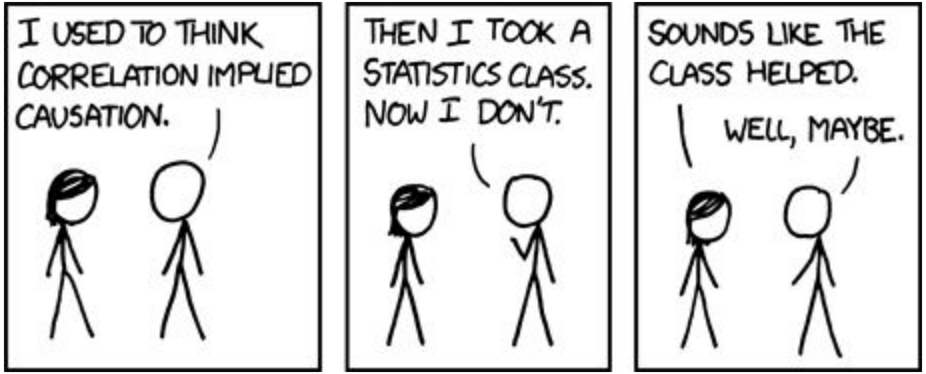
Observational causal inference in ‘Academic Self-presentation with Bibliometrics’.

And …

Perhaps attempt to answer more questions …

a. Researcher/Author level
   i. What factors (in time) influence a researcher’s change in affiliation
   ii. What factors(in time) cause a researcher’s promotion
   iii. Causes for migration/movement of researchers.
   Measuring and Understanding Migration of Scholars: Evidence from Bibliometric Data

Thank you


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.

Exploring Existing Query Datasets
(HiWi work)

Tobias Xylander

Martin-Luther-Universität Halle-Wittenberg

11.03.2022

Touche 2022 results & stances

- Roughly 25000 relevance and quality judgements
- And $\approx$ 25000 stance judgements

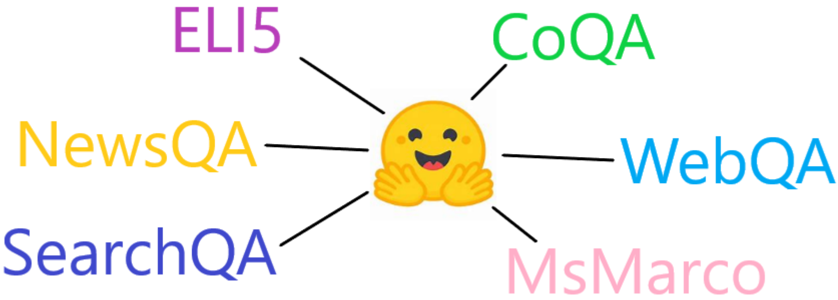
Touche 2022 results & stances

- Roughly 25000 relevance and quality judgements
- And $\approx$ 25000 stance judgements
- Of different answers to about 50 different topics
- From different annotators

Touche 2022 results & stances

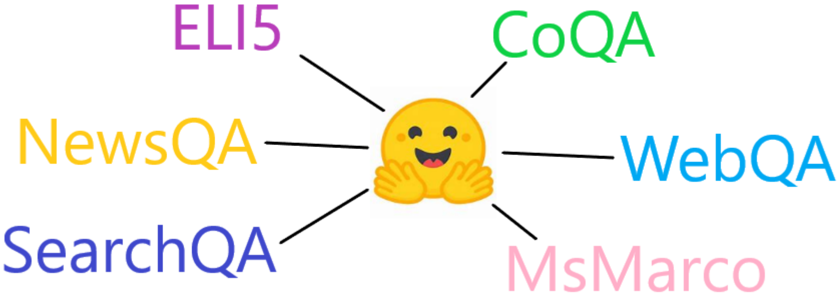
- Roughly 25000 relevance and quality judgements
- And $\approx$ 25000 stance judgements
- Of different answers to about 50 different topics
- From different annotators
- Calculated mean & majority of relevance, quality and stance for each answer

Question Answering data sets



- Download from HuggingFace

Question Answering data sets



- Download from HuggingFace
- Converting & Splitting (if necessary)
- Exploring data set statistics
- Especially data set size

Processing Multilingual MsMarco datasets

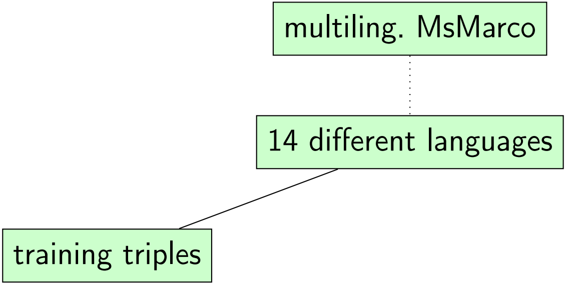
MS MARCO Passage Ranking Dataset (multilingual)

multiling. MsMarco

14 different languages

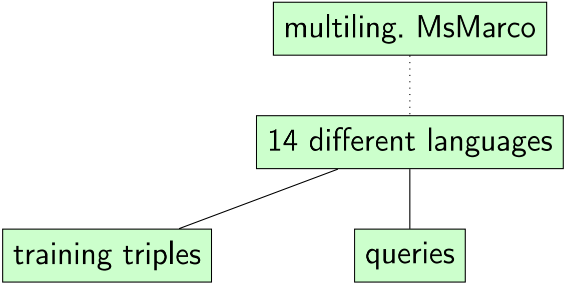
Processing Multilingual MsMarco datasets

MS MARCO Passage Ranking Dataset (multilingual)



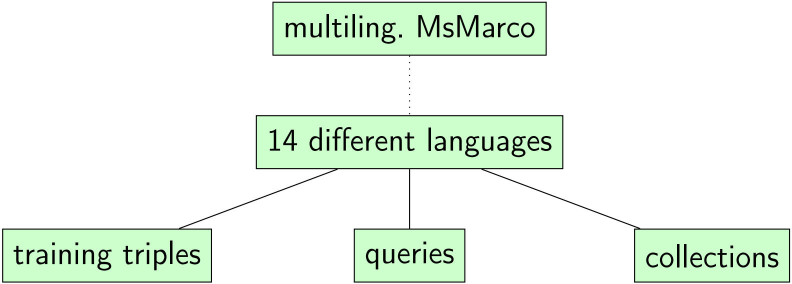
Processing Multilingual MsMarco datasets

MS MARCO Passage Ranking Dataset (multilingual)



Processing Multilingual MsMarco datasets

MS MARCO Passage Ranking Dataset (multilingual)



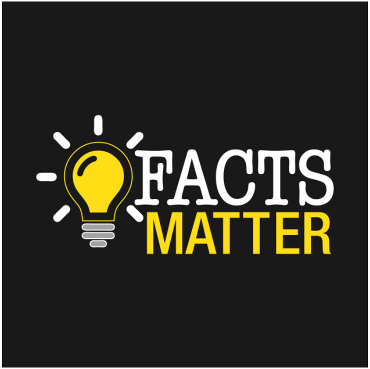
Annotating questions

Manually annotating ≈ 500 questions.
As example:" Is there any difference between faith and trust?":

Annotating questions

Manually annotating $\approx$ 500 questions.
As example:" Is there any difference between faith and trust?":

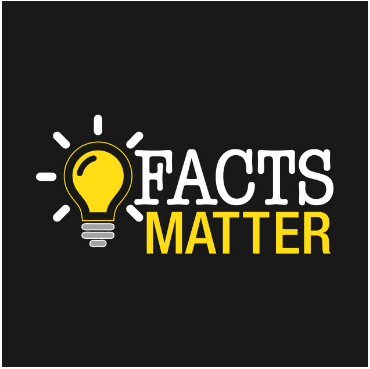
Asking for arguments?



Or only for facts?

Annotating questions

Manually annotating $\approx$ 500 questions.
As example:" Is there any difference between faith and trust?":



Asking for arguments?

Or only for facts?

Thank you.