Prompt:
Comics style. Produce an image showing a Knowledge Graphs and a Large Language Model such as GPT-4. They are used to combat COVID-19 related misinformation such as conspiracy theories.

CIMPLE

Youri Peskine
Thibault Ehrhart
Paolo Papotti
Raphaël Troncy

EURECOM
Sophia Antipolis

Hybridizing Knowledge Graph and LLM for Explaining Factors Contributing to Misinformation

raphael.troncy@eurecom.fr
@rtroncy / @CimpleXai

chist-era

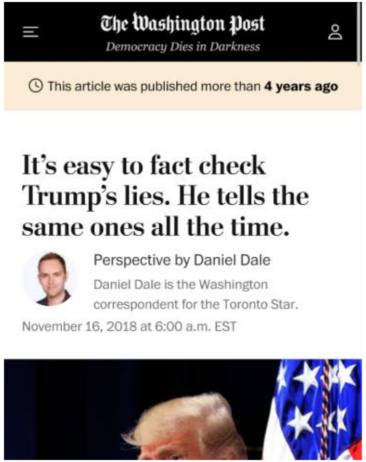
Motivation

- Online misinformation spread fast and can be dangerous [1]. According to [2], fake news spread six times faster than the corrected fact checks

- Some claims have already been fact-checked in the past. According to [3], "*viral claims often come back after a while in social media, and politicians are known to repeat the same claims over and over again.*"

- Manual fact checking is time consuming, and it does not scale well

- Current automatic approaches lack explainability [4]. Neural networks are used as "*black-boxes*"



The Washington Post
Democracy Dies in Darkness

This article was published more than **4 years ago**

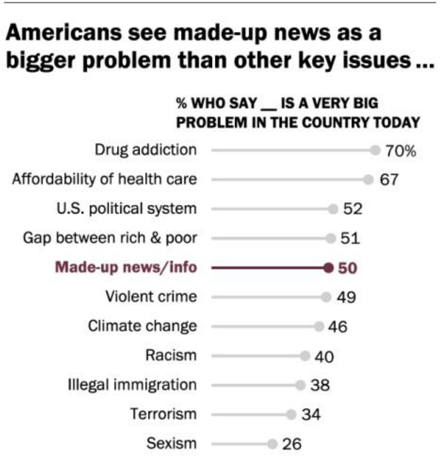
It's easy to fact check Trump's lies. He tells the same ones all the time.

Perspective by Daniel Dale

Daniel Dale is the Washington correspondent for the Toronto Star.

November 16, 2018 at 6:00 a.m. EST

[5]

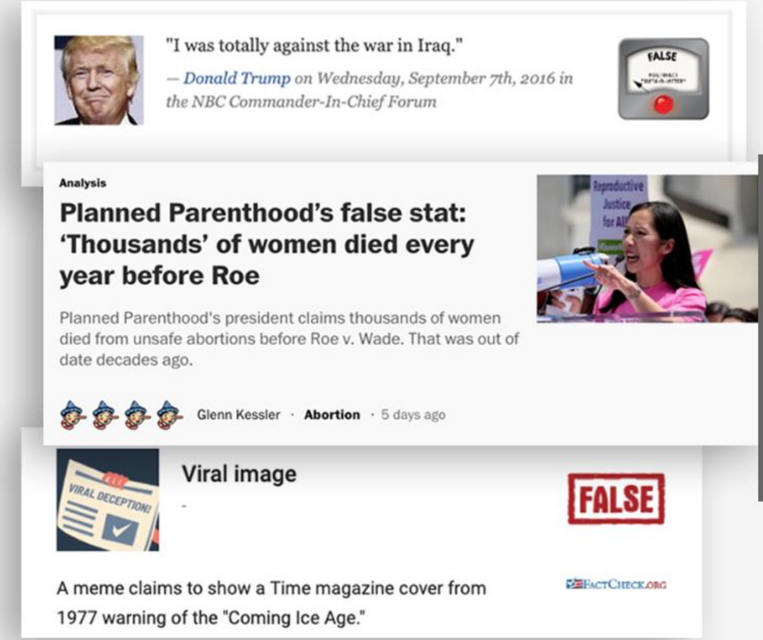


[6]

[1] https://www.who.int/health-topics/infodemic
[2] The spread of true and false news online (Vosoughi et al., Science 2018)
[3] That is a Known Lie: Detecting Previously Fact-Checked Claims (Shaar et al., ACL 2020)
[4] Towards Explainable Fact Checking (Augenstein, 2021)
[5] Washington Post news article
[6] https://www.pewresearch.org/journalism/2019/06/05/many-americans-say-made-up-news

The ClaimReview Project

English Deutsch Français Hindī Indonesian Português عربي NEW: MediaReview



"I was totally against the war in Iraq."

— *Donald Trump on Wednesday, September 7th, 2016 in the NBC Commander-In-Chief Forum*

Analysis

Planned Parenthood's false stat: 'Thousands' of women died every year before Roe

Planned Parenthood's president claims thousands of women died from unsafe abortions before Roe v. Wade. That was out of date decades ago.

Glenn Kessler · **Abortion** · 5 days ago

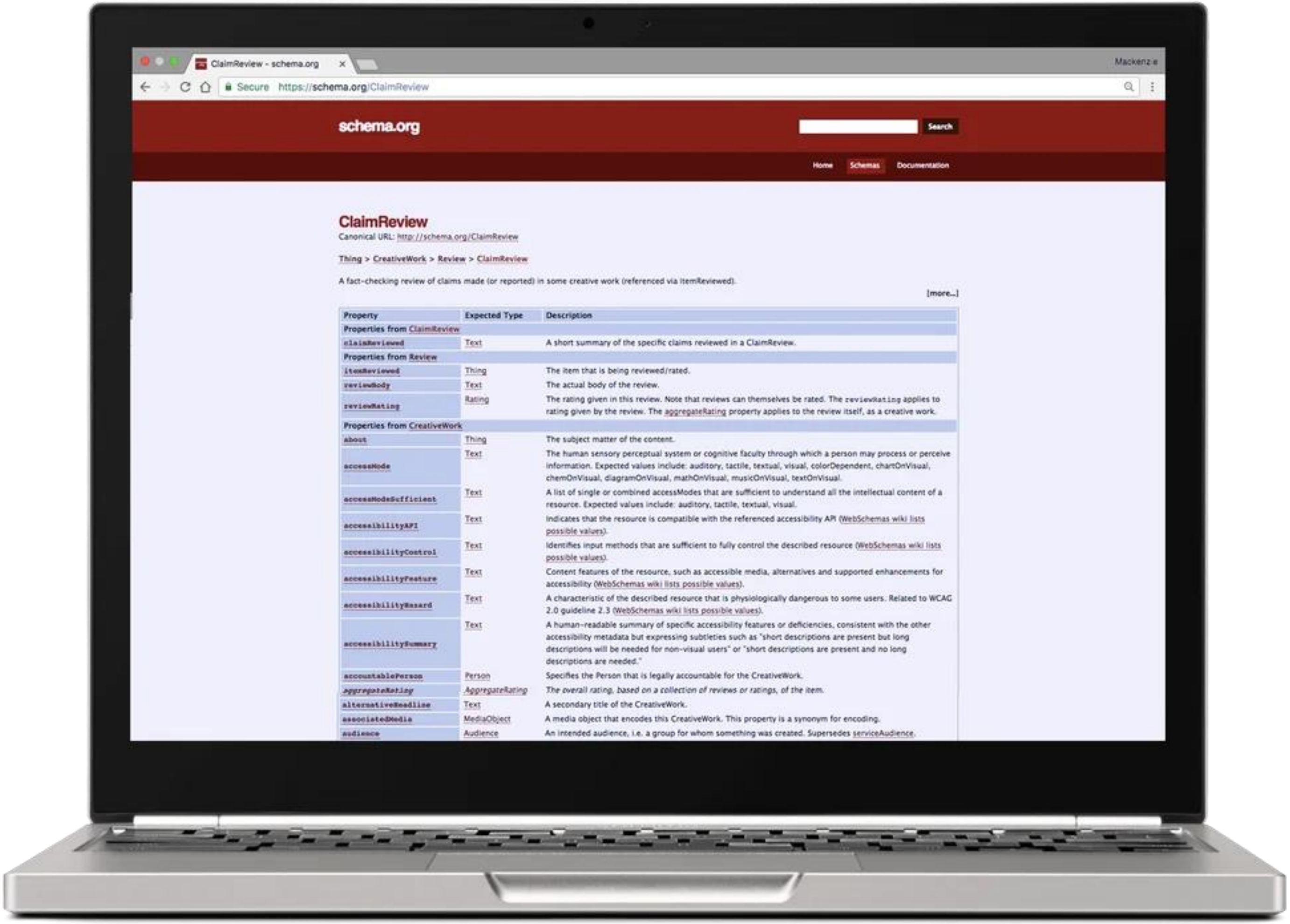
Viral image

-

A meme claims to show a Time magazine cover from 1977 warning of the "Coming Ice Age."

A tagging system for fact-checks

A new way to identify fact-check articles for search engines and apps.



ClaimReview

Canonical URL: http://schema.org/ClaimReview

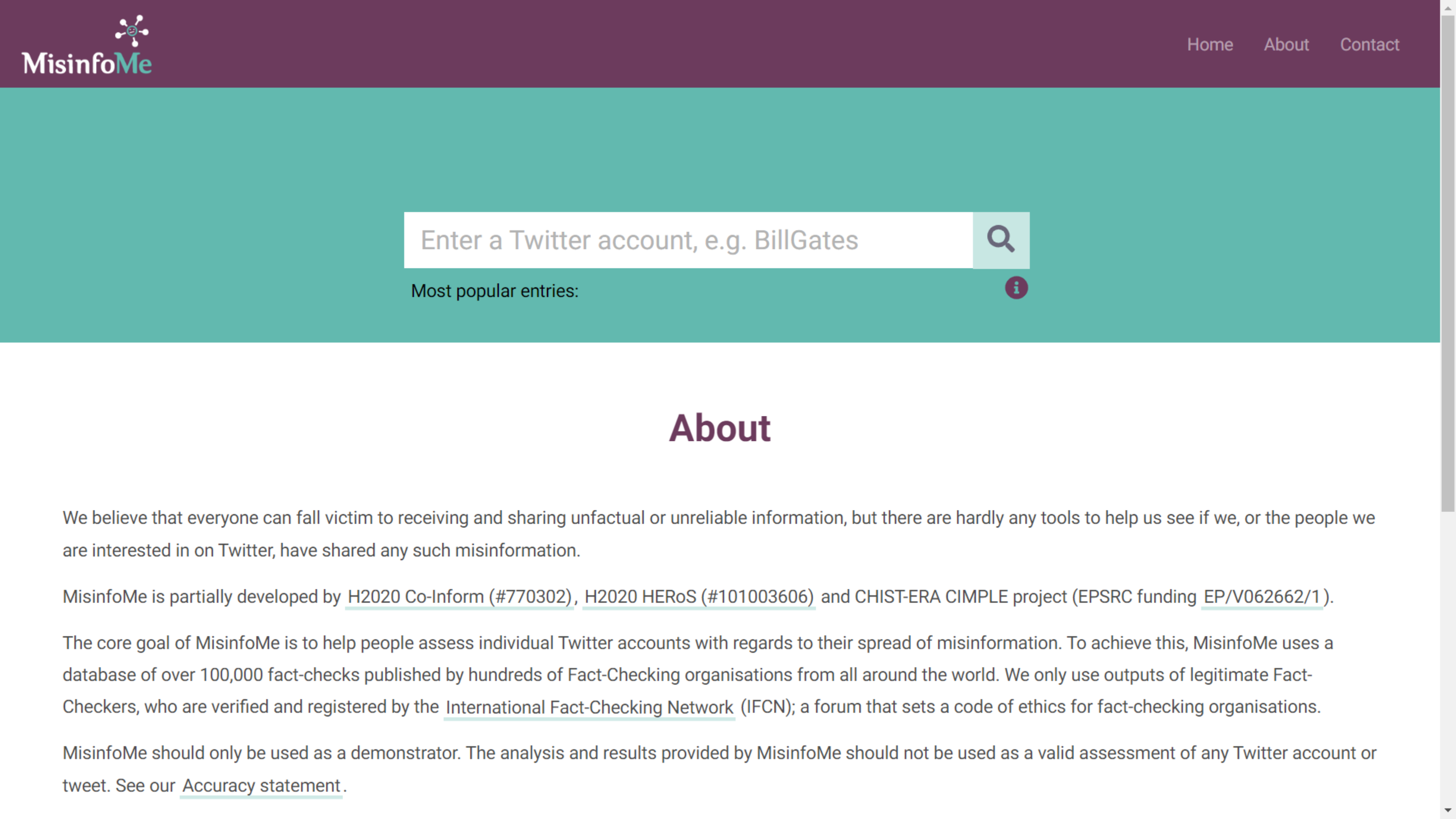
Thing > CreativeWork > Review > ClaimReview

A fact-checking review of claims made (or reported) in some creative work (referenced via itemReviewed).

[more...]

Property	Expected Type	Description
Properties from ClaimReview		
`claimReviewed`	Text	A short summary of the specific claims reviewed in a ClaimReview.
Properties from Review		
`itemReviewed`	Thing	The item that is being reviewed/rated.
`reviewBody`	Text	The actual body of the review.
`reviewRating`	Rating	The rating given in this review. Note that reviews can themselves be rated. The `reviewRating` applies to rating given by the review. The aggregateRating property applies to the review itself, as a creative work.
Properties from CreativeWork		
`about`	Thing	The subject matter of the content.
`accessMode`	Text	The human sensory perceptual system or cognitive faculty through which a person may process or perceive information. Expected values include: auditory, tactile, textual, visual, colorDependent, chartOnVisual, chemOnVisual, diagramOnVisual, mathOnVisual, musicOnVisual, textOnVisual.
`accessModeSufficient`	Text	A list of single or combined accessModes that are sufficient to understand all the intellectual content of a resource. Expected values include: auditory, tactile, textual, visual.
`accessibilityAPI`	Text	Indicates that the resource is compatible with the referenced accessibility API (WebSchemas wiki lists possible values).
`accessibilityControl`	Text	Identifies input methods that are sufficient to fully control the described resource (WebSchemas wiki lists possible values).
`accessibilityFeature`	Text	Content features of the resource, such as accessible media, alternatives and supported enhancements for accessibility (WebSchemas wiki lists possible values).
`accessibilityHazard`	Text	A characteristic of the described resource that is physiologically dangerous to some users. Related to WCAG 2.0 guideline 2.3 (WebSchemas wiki lists possible values).
`accessibilitySummary`	Text	A human-readable summary of specific accessibility features or deficiencies, consistent with the other accessibility metadata but expressing subtleties such as "short descriptions are present but long descriptions will be needed for non-visual users" or "short descriptions are present and no long descriptions are needed."
`accountablePerson`	Person	Specifies the Person that is legally accountable for the CreativeWork.
`aggregateRating`	*AggregateRating*	*The overall rating, based on a collection of reviews or ratings, of the item.*
`alternativeHeadline`	Text	A secondary title of the CreativeWork.
`associatedMedia`	MediaObject	A media object that encodes this CreativeWork. This property is a synonym for encoding.
`audience`	Audience	An intended audience, i.e. a group for whom something was created. Supersedes serviceAudience.

https://blog.google/products/search/fact-check-now-available-google-search-and-news-around-world/



About

We believe that everyone can fall victim to receiving and sharing unfactual or unreliable information, but there are hardly any tools to help us see if we, or the people we are interested in on Twitter, have shared any such misinformation.

MisinfoMe is partially developed by H2020 Co-Inform (#770302), H2020 HERoS (#101003606) and CHIST-ERA CIMPLE project (EPSRC funding EP/V062662/1).

The core goal of MisinfoMe is to help people assess individual Twitter accounts with regards to their spread of misinformation. To achieve this, MisinfoMe uses a database of over 100,000 fact-checks published by hundreds of Fact-Checking organisations from all around the world. We only use outputs of legitimate Fact-Checkers, who are verified and registered by the International Fact-Checking Network (IFCN); a forum that sets a code of ethics for fact-checking organisations.

MisinfoMe should only be used as a demonstrator. The analysis and results provided by MisinfoMe should not be used as a valid assessment of any Twitter account or tweet. See our Accuracy statement.

Detection of factors that could explain misinformation

Analyzing COVID-Related Social Discourse on Twitter using Emotion, Sentiment, Political Bias, Stance, Veracity and Conspiracy Theories

DOI: https://dl.acm.org/doi/10.1145/3543873.3587622

Best paper award at the BeyondFacts Workshop,
colocated with The Web Conference 2023, Austin, USA

THE WEB CONFERENCE ACM

COVID-19 misinformation

Many datasets have been built around annotating textual features in tweets during the pandemic

- COVID LTSE Attributes [1]
- COVIDSenti [2]
- Russian Troll [3]
- COVID 19 Stance [4]
- Birdwatch [5]
- MediaEval FND [6]

[1] Gupta, R., Vishwanath, A., & Yang, Y.. (2022). **COVID-19 Twitter Dataset with Latent Topics, Sentiments and Emotions Attributes**.
[2] Naseem, U., Razzak, I., Khushi, M., Eklund, P., & Kim, J. (2021). **COVIDSenti: A large-scale benchmark Twitter data set for COVID-19 sentiment analysis.**
[3] Darren L. Linvill, & Patrick L. Warren (2020). **Troll Factories: Manufacturing Specialized Disinformation on Twitter**.
[4] Glandt, K., Khanal, S., Li, Y., Caragea, D., & Caragea, C. (2021). **Stance Detection in COVID-19 Tweets.**
[5] Saeed, M., Traub, N., Nicolas, M., Demartini, G., & Papotti, P. (2022). **Crowdsourced Fact-Checking at Twitter: How Does the Crowd Compare With Experts?**
[6] Konstantin Pogorelov, Daniel Thilo Schroeder, Stefan Brenner, and Asep Maulana, & Johannes Langguth (2022). **Combining Tweets and Connections Graph for FakeNews Detection at MediaEval 2022**. See also https://link.springer.com/article/10.1007/s42001-023-00200-3

Tweet Datasets

All dataset contain tweets. All are about COVID (except Russian Troll)

Dataset	Samples	Start date	End date	Labels
COVID LTSE Attributes	252M	January 2020	June 2021	Emotion (*fear, anger, joy, sadness, no emotion*)
COVIDSenti	90,000	February 2020	March 2020	Sentiment (*positive, negative, neutral*)
Russian Troll	2.9M	February 2012	May 2018	Political Leaning (*right, left, other*)
COVID 19 Stance	3,616	February 2020	August 2020	Stance (*in-favor, against*)
Birdwatch	9,851	January 2021	September 2021	Veracity (*misleading, not misleading*)
MediaEval FND (COCO)	1,912	January 2020	June 2021	Conspiracy Theories (*9 named conspiracy theories*)

Factors that affect Misinformation

Factors are textual features that allow a better understanding of documents

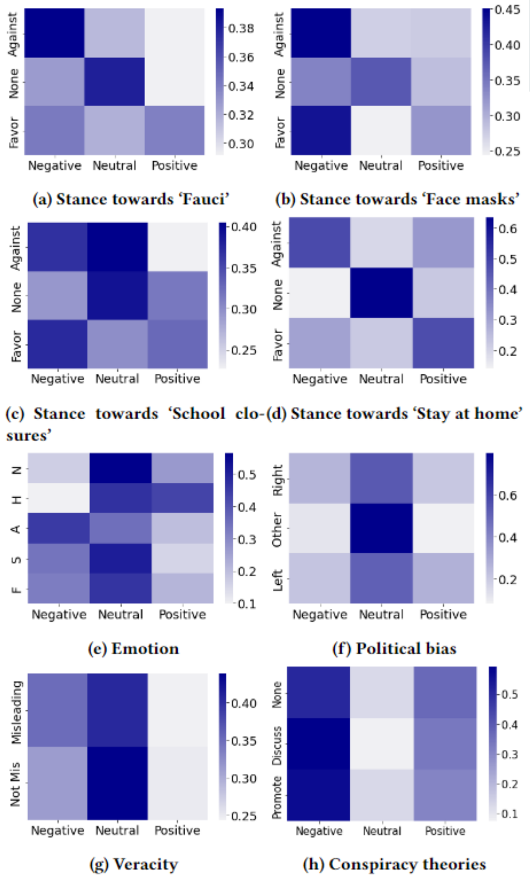
Train BERT-based models$^{(*)}$ to detect:

- Emotion, Sentiment, Political-leaning [2]
- Conspiracy Theories [1]
- Propaganda Techniques (ongoing)

Model	F1-score
Emotion	0.622
Sentiment	0.769
Political-bias	0.636

$^{(*)}$ CT-BERT models with weighted Cross Entropy Loss

[1] https://github.com/D2KLab/mediaeval-fakenews
[2] https://github.com/D2KLab/covid-twitter-discourse-analysis



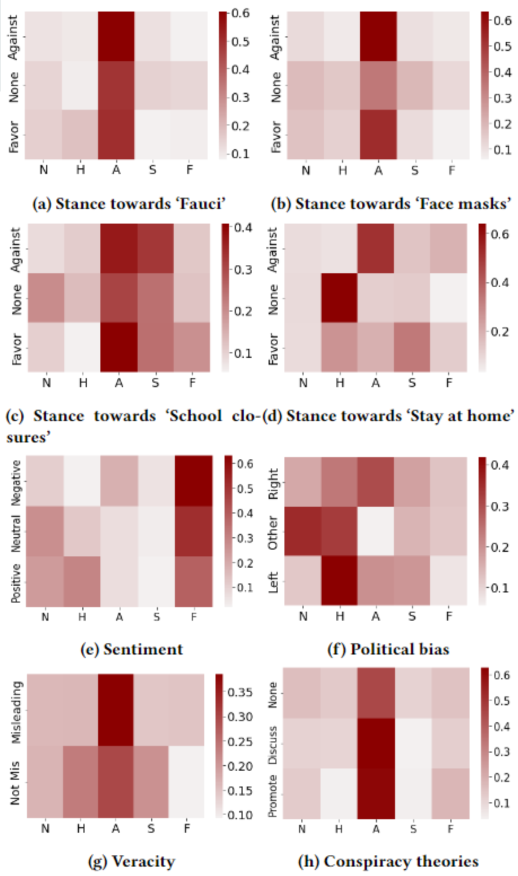
(a) Stance towards 'Fauci' (b) Stance towards 'Face masks'

(c) Stance towards 'School closures' (d) Stance towards 'Stay at home'

(e) Emotion (f) Political bias

(g) Veracity (h) Conspiracy theories

Figure 1: Distribution of the labels for the sentiment feature



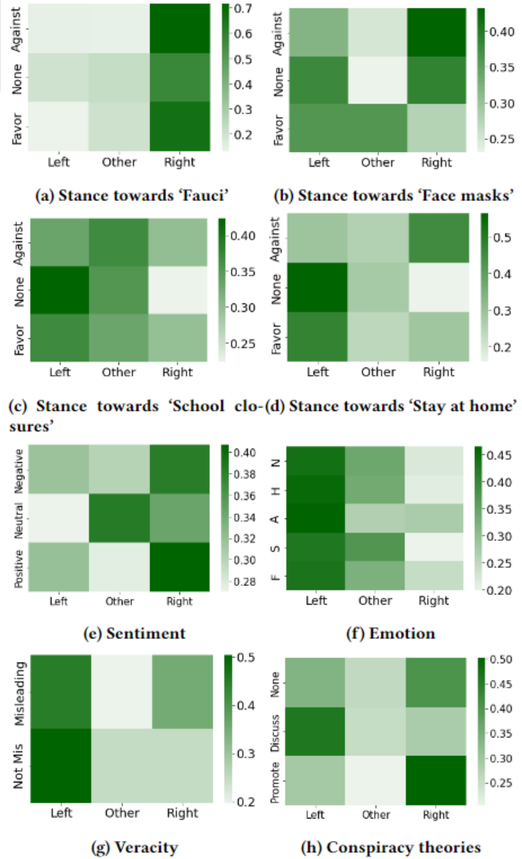
(a) Stance towards 'Fauci' (b) Stance towards 'Face masks'

(c) Stance towards 'School closures' (d) Stance towards 'Stay at home'

(e) Sentiment (f) Political bias

(g) Veracity (h) Conspiracy theories

Figure 2: Distribution of the labels for the emotion feature



(a) Stance towards 'Fauci' (b) Stance towards 'Face masks'

(c) Stance towards 'School closures' (d) Stance towards 'Stay at home'

(e) Sentiment (f) Emotion

(g) Veracity (h) Conspiracy theories

Figure 3: Distribution of the labels for the political bias feature

TL;DR: the users' political leaning reflect the stance of US politicians in the debate, conspiracy theories are usually promoted with negative sentiment and right political leaning, COVID topics are highly controversial on Twitter

Detection of COVID-19-related conspiracy theories in tweets using transformer-based models and node embedding techniques

Multimedia Evaluation Workshop (MediaEval), FakeNews Detection Task,
12-13 January 2023, Bergen, Norway
https://2022.multimediaeval.com/paper8669.pdf

MediaEval Benchmark
MediaEval Benchmarking Initiative for Multimedia Evaluation
The "multi" in multimedia: speech, audio, visual content, tags, users, context

MediaEval FakeNews 2021 and 2022

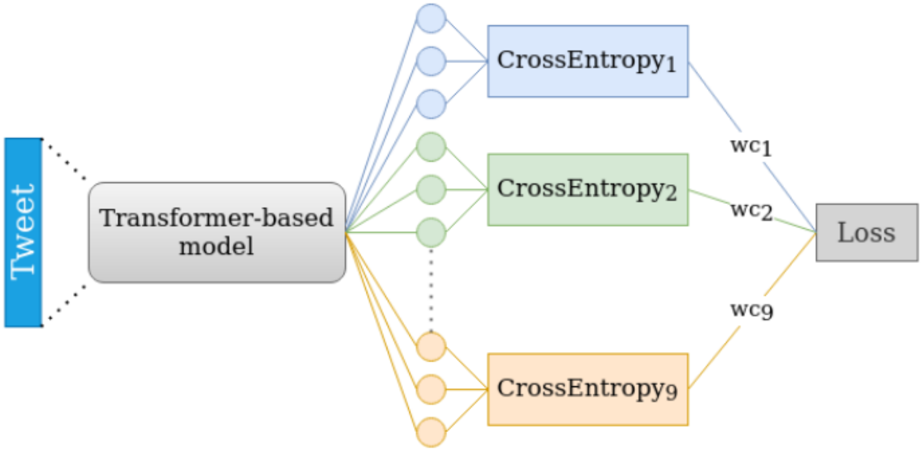
Goal: Detect 9 different named COVID- related conspiracy theories* in tweets.

“*I took a container of lysol wipes and looked at the back and it indeed does say helps with human coronavirus. How can it be a new disease when it's been on lysol containers for a long time? It isnt new,that's how. Someone is using it for biological warfare? Population control?*”

(Supporting Intentional pandemic, Discussing Population reduction)

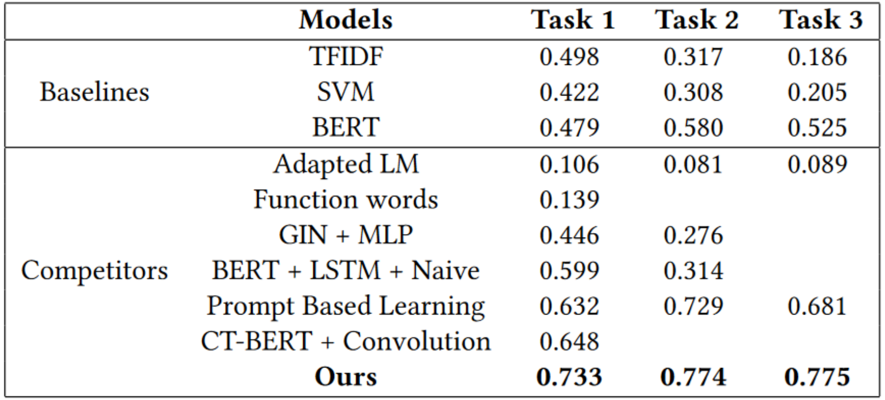
* 9 theories: Suppressed Cures, Behaviour and Mind Control, Antivax, Fake virus, Intentional Pandemic, Harmful Radiation/ Influence, Population reduction, New World Order, Satanism

Approach and Results



Approach: Transformer-based architecture (RoBERTa, CT-bert)

Table 1: MCC results for baselines and competitors on the test set of the ME21FND challenge



	Models	Task 1	Task 2	Task 3
Baselines	TFIDF	0.498	0.317	0.186
	SVM	0.422	0.308	0.205
	BERT	0.479	0.580	0.525
Competitors	Adapted LM	0.106	0.081	0.089
	Function words	0.139		
	GIN + MLP	0.446	0.276	
	BERT + LSTM + Naive	0.599	0.314	
	Prompt Based Learning	0.632	0.729	0.681
	CT-BERT + Convolution	0.648		
	Ours	**0.733**	**0.774**	**0.775**

1st place in all 3 tasks

distinctive mention

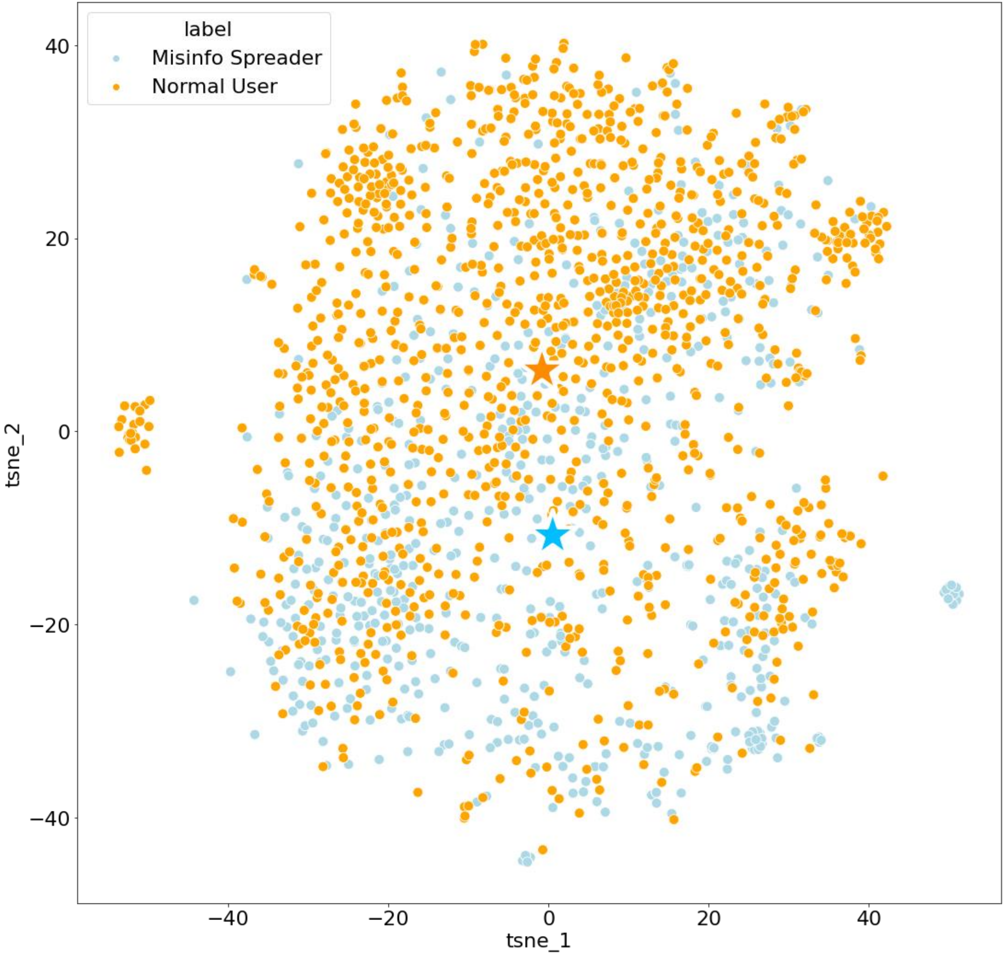
Detecting Misinformation Spreader

Using the user-interaction graph

1. Build the directed graph from the user graph (1.7M nodes, 270M edges)
2. Generate random walks on the graph (r, l, p, q)
3. Train word2vec model on the random walks (32 dimensions)
4. Train different ML algorithms on embeddings

t-sne algorithm to visualize data in two dimension

Plot corresponds to the 2k users provided in the training data



Classification Results: is a user a Misinformation Spreader?

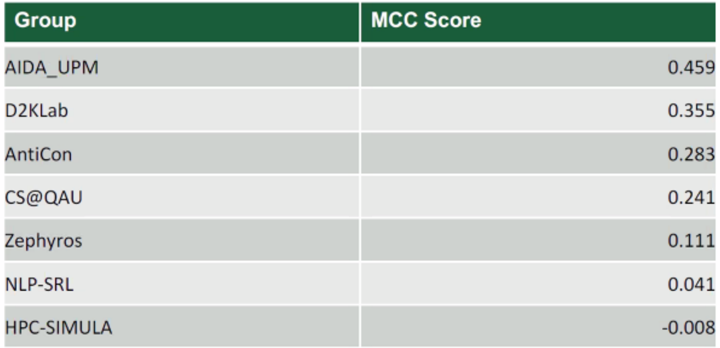
Which graph embeddings algorithm to use?

- *node2vec* uses the structure of the graph and the neighbors of a node to learn embeddings
- Train different machine learning algorithms from the 32 features
- MLP layers: 32 - 16 - 8 – 1
- Default parameters from sklearn

Model	Mean MCC
MLP	**0.447**
SVC	0.398
KNeighborsClassifier	0.353
GaussianProcessClassifier	0.320
DecisionTreeClassifier	0.315
RandomForestClassifier	**0.428**
AdaBoostClassifier	0.342
GaussianNB	0.361
QuadraticDiscriminantAnalysis	0.287
GradientBoostingClassifier	0.406

What about GNNs? … they all failed!

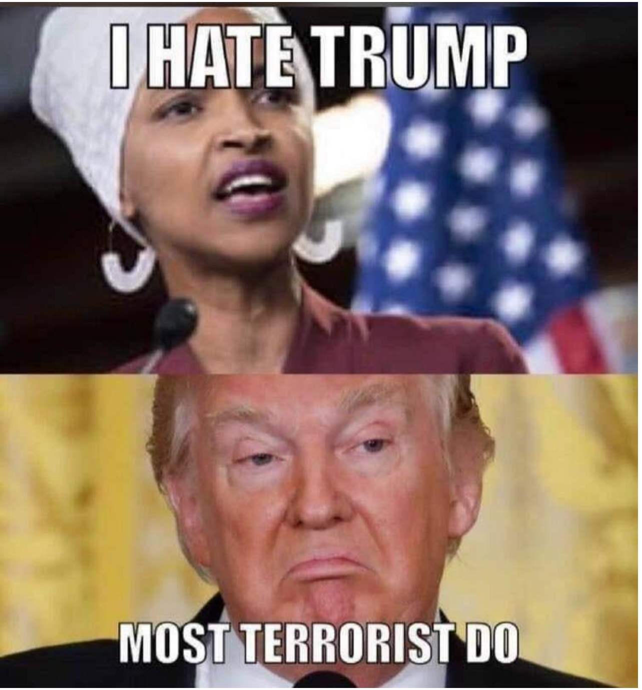
… more on this later

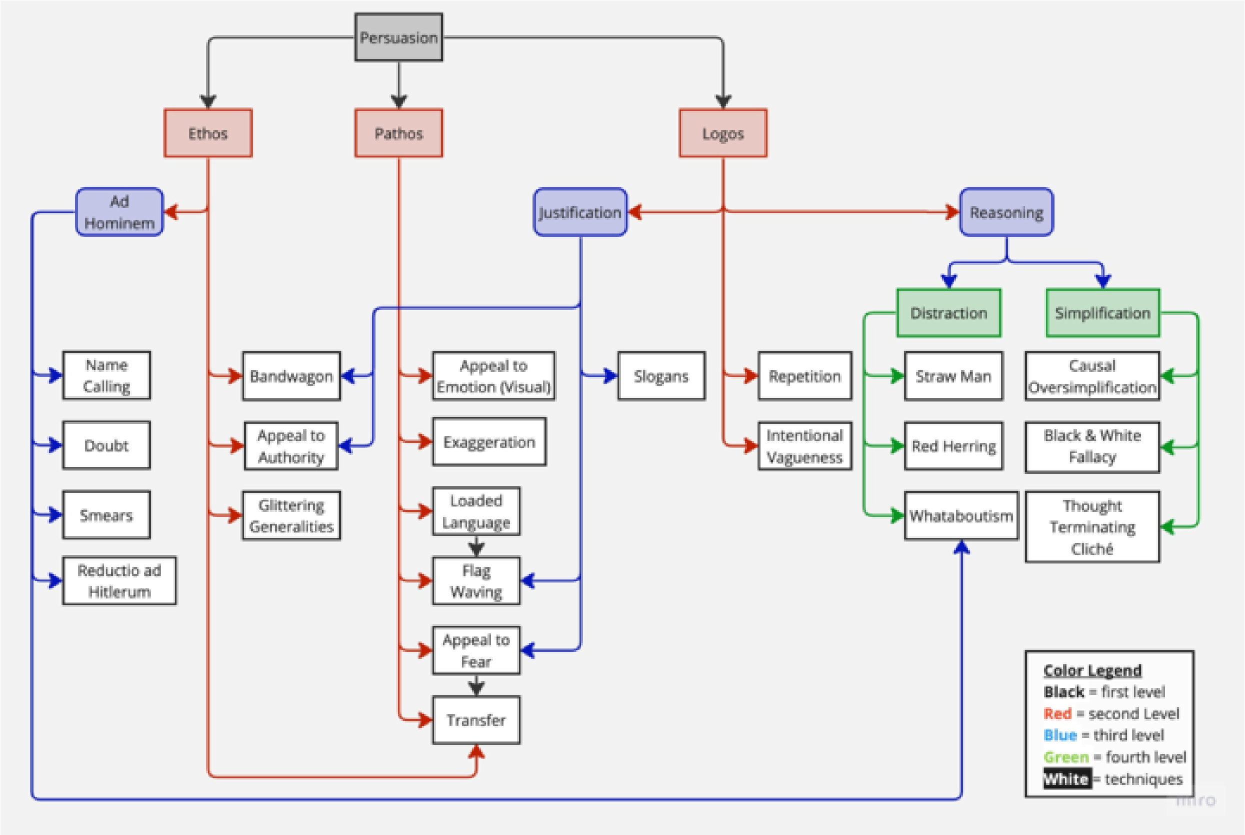


Group	MCC Score
AIDA_UPM	0.459
D2KLab	0.355
AntiCon	0.283
CS@QAU	0.241
Zephyros	0.111
NLP-SRL	0.041
HPC-SIMULA	-0.008

SemEval 2024 - Task 4

Detection of Persuasion Techniques in Memes (ongoing work)





Persuasion

Ethos

Pathos

Logos

Ad Hominem

Justification

Reasoning

Distraction

Simplification

Name Calling

Doubt

Smears

Reductio ad Hitlerum

Bandwagon

Appeal to Authority

Glittering Generalities

Appeal to Emotion (Visual)

Exaggeration

Loaded Language

Flag Waving

Appeal to Fear

Transfer

Slogans

Repetition

Intentional Vagueness

Straw Man

Red Herring

Whataboutism

Causal Oversimplification

Black & White Fallacy

Thought Terminating Cliché

Color Legend

Black = first level

Red = second Level

Blue = third level

Green = fourth level

White = techniques

Definitions Matter: Guiding GPT for Multi-label Classification

International Conference on Empirical Methods in Natural Language Processing (EMNLP Findings)
6-10 December 2023, Singapore

EMNLP 2023

Text Classification using LLM

Analyze performance of GPT-3 model for fine-grained multi-label conspiracy theory classification of tweets using zero-shot labeling.

Three different experiments, changing the prompt to include:

1. The textual label
2. The textual label and a generated definition (EG)
3. The textual label and a human-written definition (HW)

Contributions:

1. Generating definitions from examples with a LLM
2. Experiments to evaluate GPT-3's understanding of the definitions.

Generating definitions

SYSTEM = "You will be given two sets of tweets. The first set of tweets contains examples of texts that mention the same concept. The second set of tweets contains examples of texts that mention other concepts, but not the same concept that tweets from the first set. Your task is to provide the definition of the concept present in the first set"

USER = "First set of tweets:

[25x Tweets containing the conspiracy]

Second set of tweets:

[25x Tweets not containing the conspiracy]

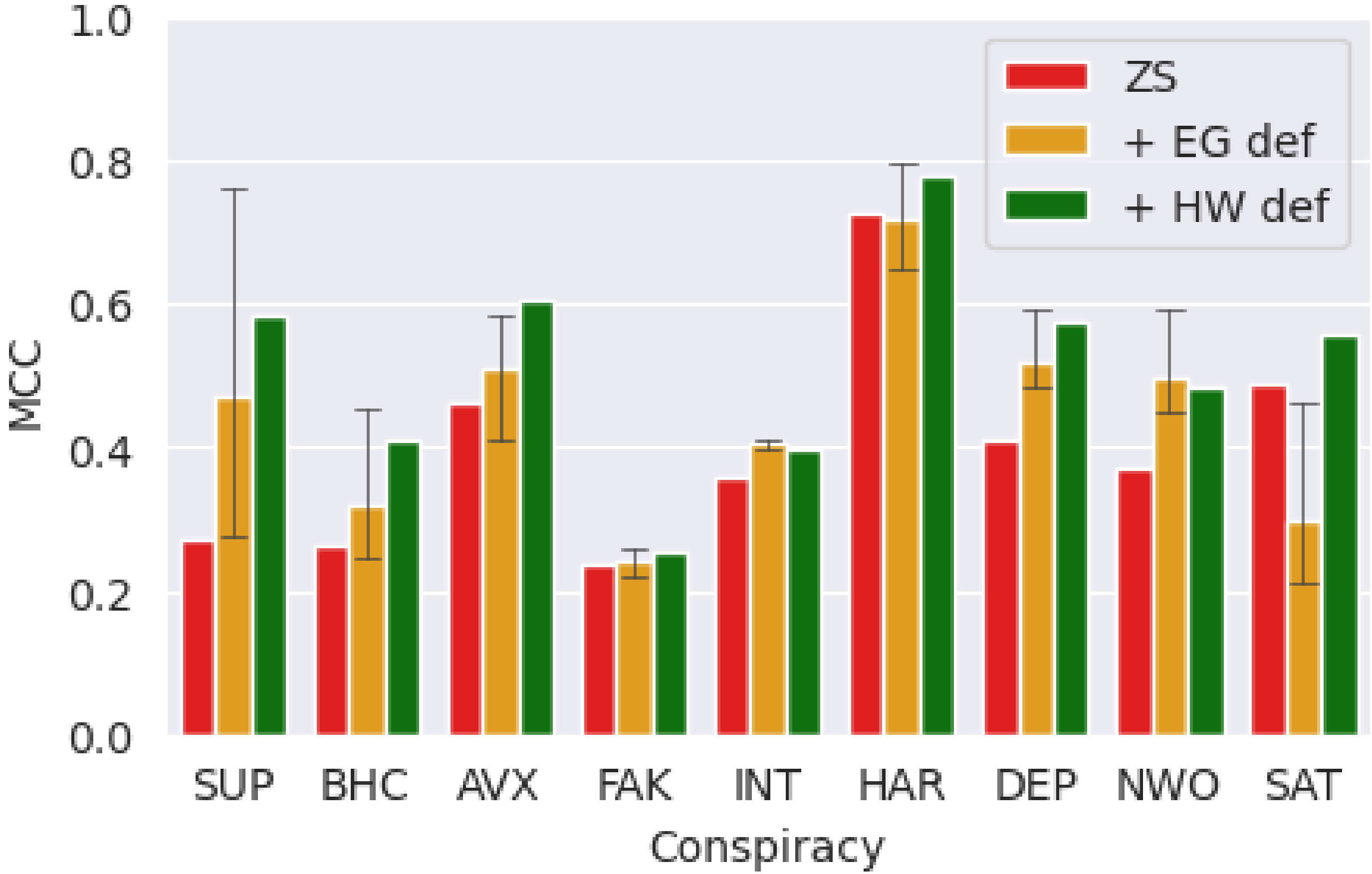
Given those two sets of tweets, what is the definition of the concept present in the first set that is not present in the second set of tweets? Start your answer with: 'The definition of the concept is' "

GPT-3: **The definition of the concept is a conspiracy theory that claims that the COVID-19 pandemic is not caused by a virus, but rather by the implementation of 5G technology and that the pandemic is being used as a distraction to implement a digital currency and implantable microchips for population control. This theory has been widely discredited by scientific evidence and health organizations.**

ChatGPT experiments on MediaEval: MCC results

Supervised approach:

0.717
MCC



Definition understanding

Better definitions lead to better performance

GPT-3 is able to correctly interpret the definitions and apply them to conspiracy theory detection

Corr(Similarity (EG, HW), MCC) = 0.375

Corr(Similarity (EG, HW), F1) = 0.390

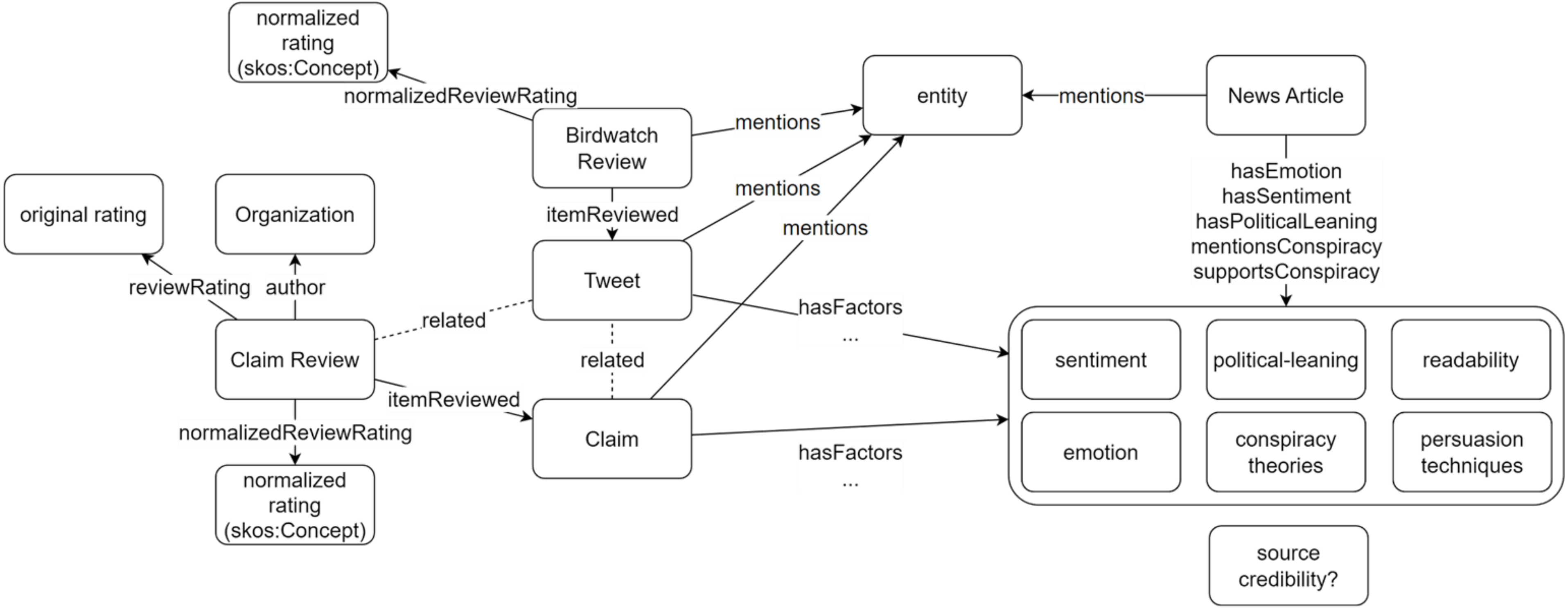
Corr(Similarity (EG$_1$, EG$_2$), Cohen’s κ) = 0.407

Corr(length(EG), MCC) = 0.062

Putting all together:
Meet the CIMPLE Knowledge Graph

CIMPLE KG - Schema

Ontology, Controlled vocabularies and converter: https://github.com/CIMPLE-project/converter
Nightly update: https://github.com/MartinoMensio/claimreview-data

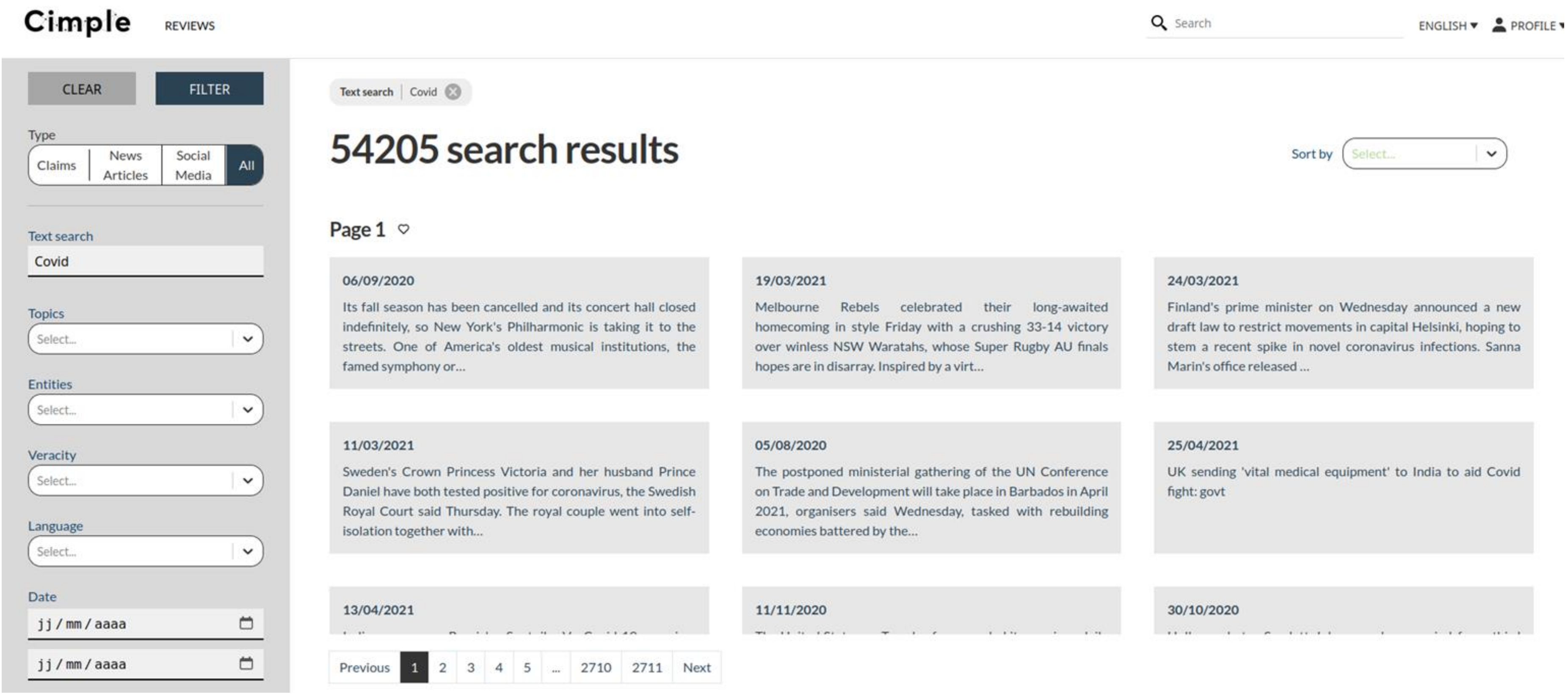


CIMPLE KG - Dataset Statistics

Filter: [January 2020 - June 2021]. SPARQL endpoint: https://data.cimple.eu/sparql

Datasets	# Documents	Update
AFP	193,933 News Articles	Snapshot
Claim Review	17,947 Claims 17,996 ClaimReview	Live (nightly update)
Birdwatch *Community Notes*	6,563 Tweets + 1,983 Reviews 736 links to ClaimReview	Snapshot *Live?*
Check-That (2022)	1,196 Tweets 16 links to ClaimReview (+12 new ones)	Snapshot
COCO (MediaEval (2021-2022)	2,702 Tweets	Snapshot
PTC (Propaganda SemEval 2020)	1,908 Claims	Snapshot

Demo at https://explorer.cimple.eu/



Text search: Covid

Text search | Covid

54205 search results

Page 1

06/09/2020
Its fall season has been cancelled and its concert hall closed indefinitely, so New York's Philharmonic is taking it to the streets. One of America's oldest musical institutions, the famed symphony or...

19/03/2021
Melbourne Rebels celebrated their long-awaited homecoming in style Friday with a crushing 33-14 victory over winless NSW Waratahs, whose Super Rugby AU finals hopes are in disarray. Inspired by a virt...

24/03/2021
Finland's prime minister on Wednesday announced a new draft law to restrict movements in capital Helsinki, hoping to stem a recent spike in novel coronavirus infections. Sanna Marin's office released ...

11/03/2021
Sweden's Crown Princess Victoria and her husband Prince Daniel have both tested positive for coronavirus, the Swedish Royal Court said Thursday. The royal couple went into self-isolation together with...

05/08/2020
The postponed ministerial gathering of the UN Conference on Trade and Development will take place in Barbados in April 2021, organisers said Wednesday, tasked with rebuilding economies battered by the...

25/04/2021
UK sending 'vital medical equipment' to India to aid Covid fight: govt

13/04/2021

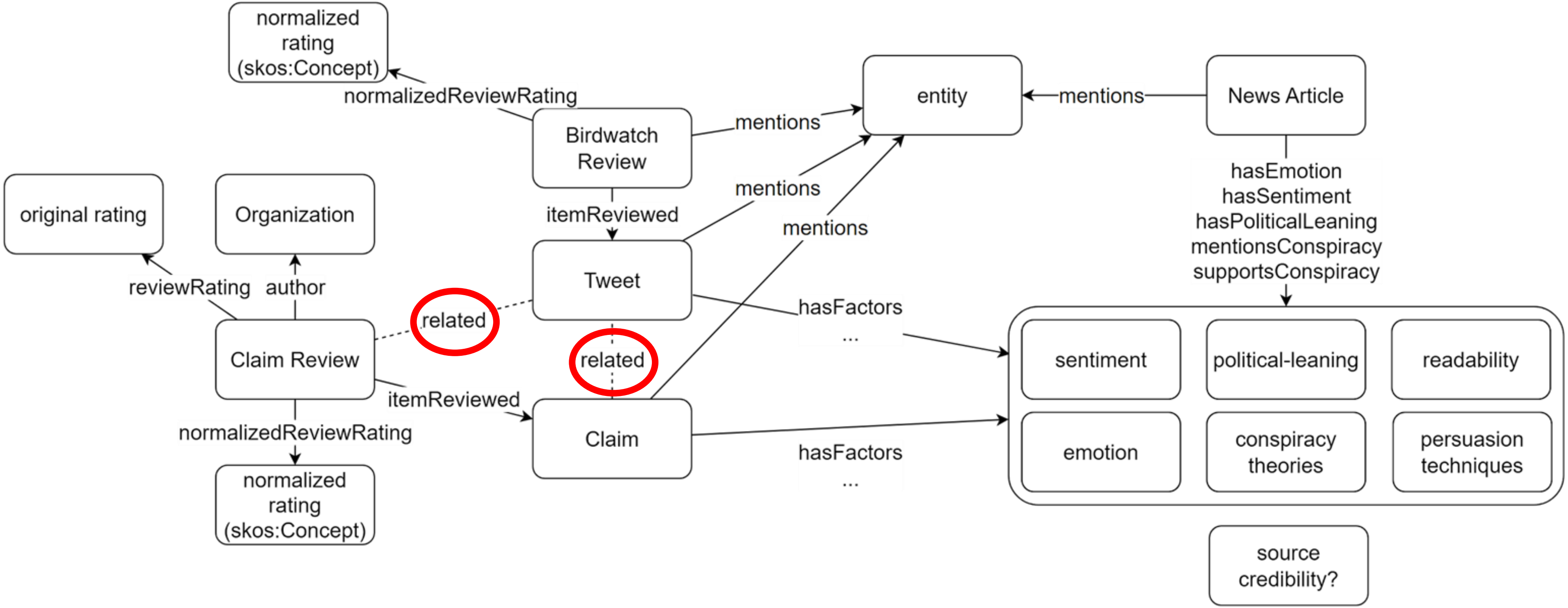
11/11/2020

30/10/2020

Example: https://explorer.cimple.eu/claim-review/dc85c1242d6edbf7043a929d6ac34969f76643e7ba76a7396554fa40

CIMPLE KG - Schema

Ontology, Controlled vocabularies and converter: https://github.com/CIMPLE-project/converter



Text Similarity

AFP1

Macron makes 'end of summer' vaccine pledge to France
Emmanuel Macron said on Tuesday that all of his countrymen who wanted a vaccine would be offered one "by the end of the summer".'

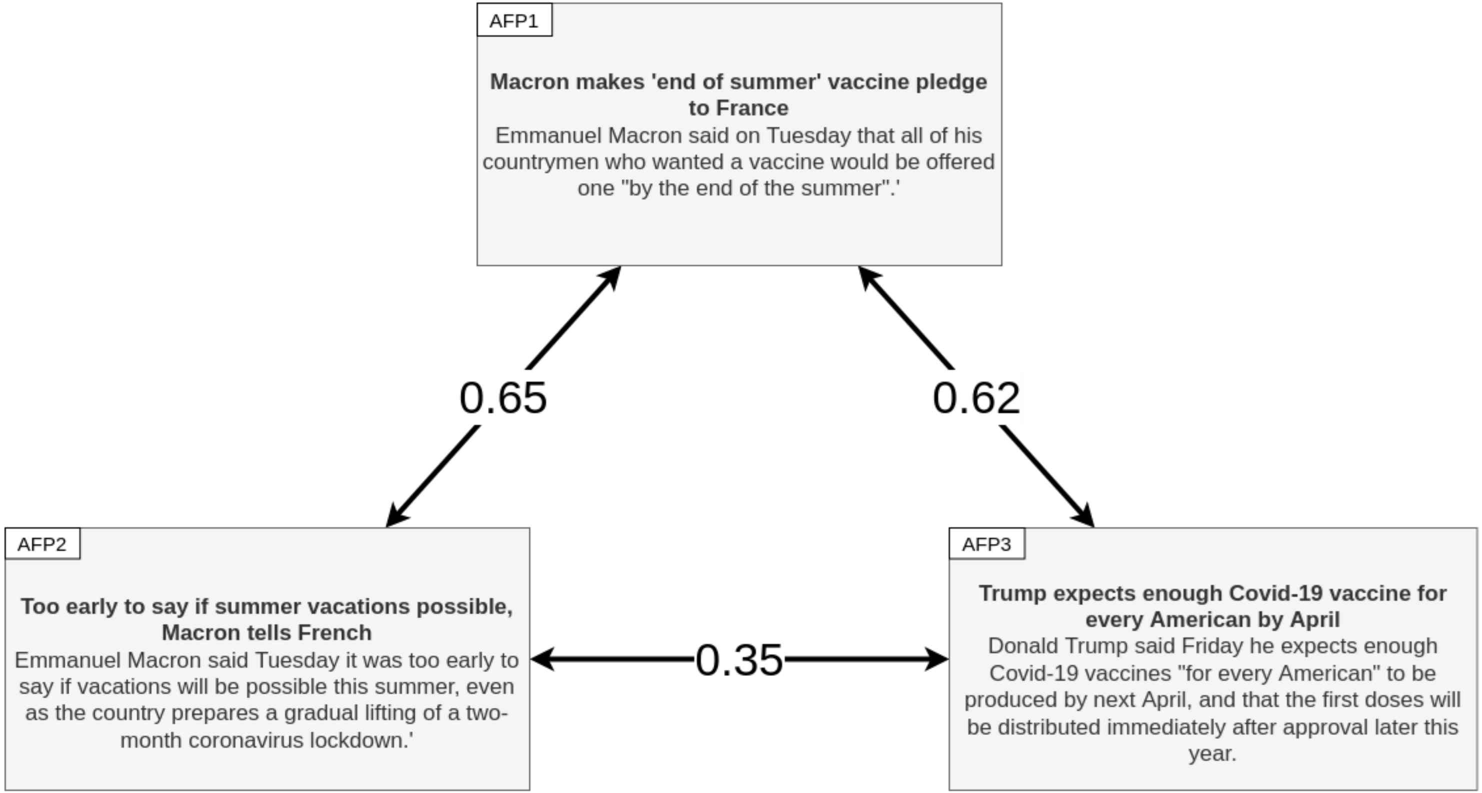
AFP2

Too early to say if summer vacations possible, Macron tells French
Emmanuel Macron said Tuesday it was too early to say if vacations will be possible this summer, even as the country prepares a gradual lifting of a two-month coronavirus lockdown.'

AFP3

Trump expects enough Covid-19 vaccine for every American by April
Donald Trump said Friday he expects enough Covid-19 vaccines "for every American" to be produced by next April, and that the first doses will be distributed immediately after approval later this year.

Text Similarity: using SentenceBERT

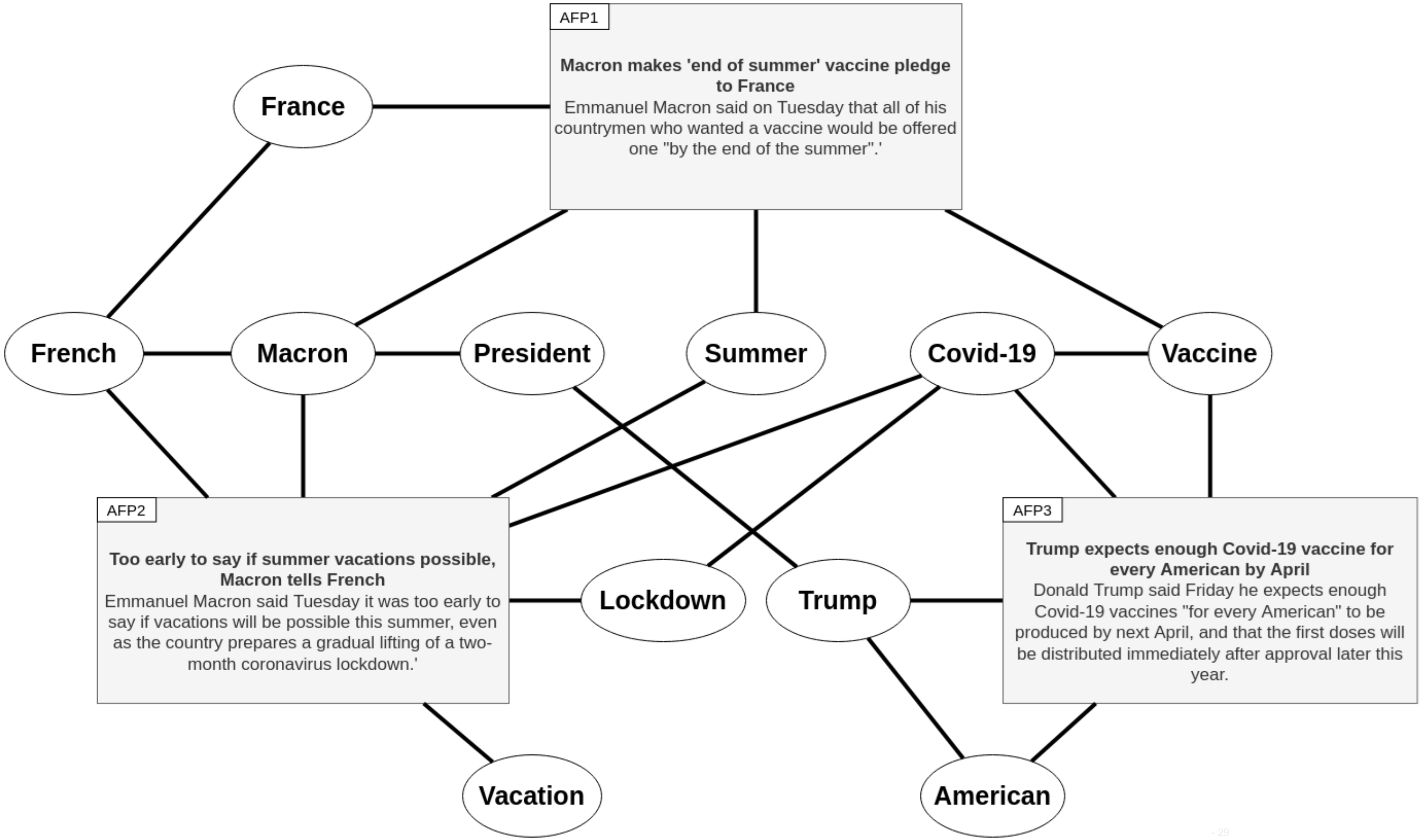


Macron makes 'end of summer' vaccine pledge to France
Emmanuel Macron said on Tuesday that all of his countrymen who wanted a vaccine would be offered one "by the end of the summer".'

Too early to say if summer vacations possible, Macron tells French
Emmanuel Macron said Tuesday it was too early to say if vacations will be possible this summer, even as the country prepares a gradual lifting of a two-month coronavirus lockdown.'

Trump expects enough Covid-19 vaccine for every American by April
Donald Trump said Friday he expects enough Covid-19 vaccines "for every American" to be produced by next April, and that the first doses will be distributed immediately after approval later this year.

Text Similarity: using entities



AFP1

Macron makes 'end of summer' vaccine pledge to France
Emmanuel Macron said on Tuesday that all of his countrymen who wanted a vaccine would be offered one "by the end of the summer".'

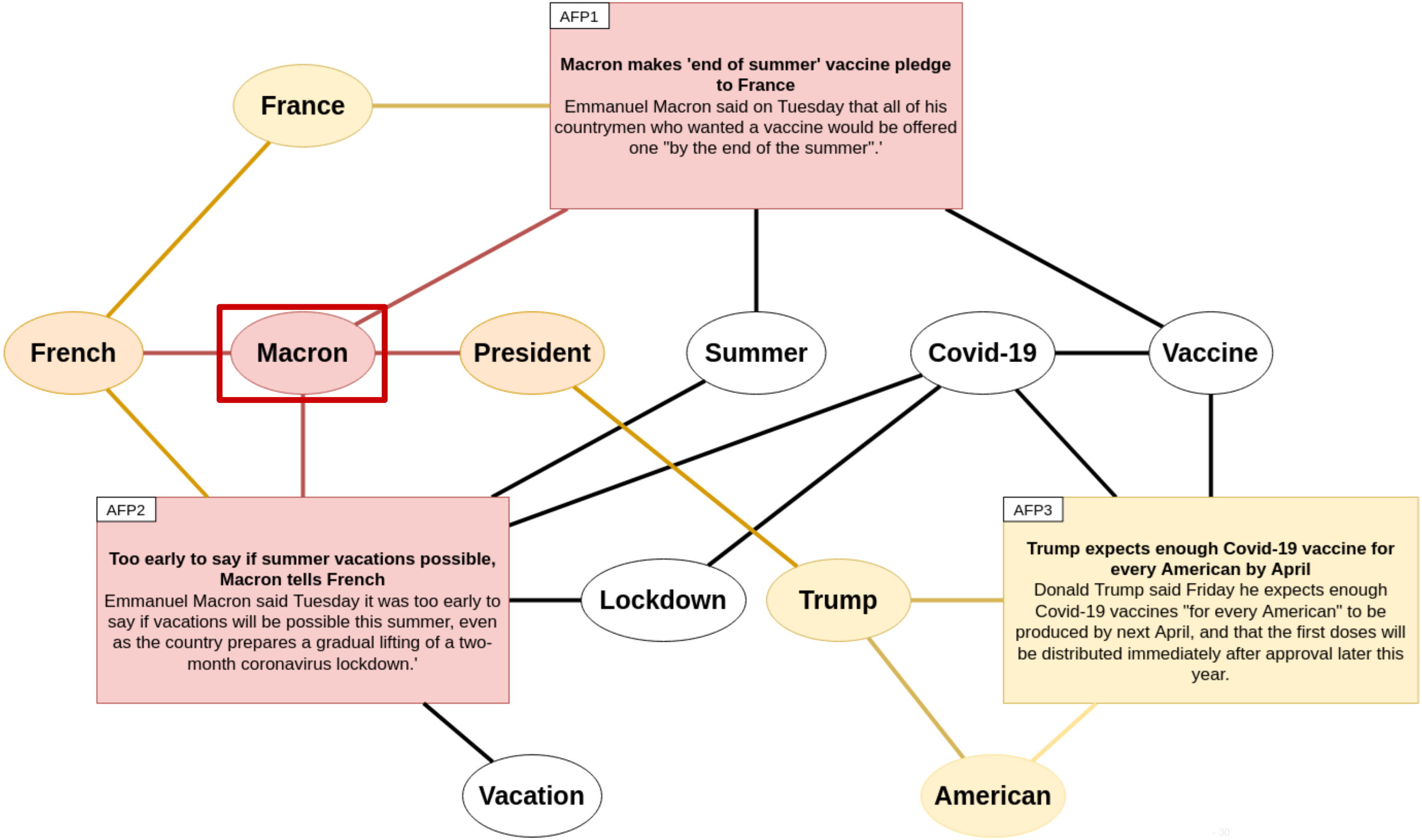
AFP2

Too early to say if summer vacations possible, Macron tells French
Emmanuel Macron said Tuesday it was too early to say if vacations will be possible this summer, even as the country prepares a gradual lifting of a two-month coronavirus lockdown.'

AFP3

Trump expects enough Covid-19 vaccine for every American by April
Donald Trump said Friday he expects enough Covid-19 vaccines "for every American" to be produced by next April, and that the first doses will be distributed immediately after approval later this year.

Text Similarity: using entities



AFP1

Macron makes 'end of summer' vaccine pledge to France
Emmanuel Macron said on Tuesday that all of his countrymen who wanted a vaccine would be offered one "by the end of the summer".'

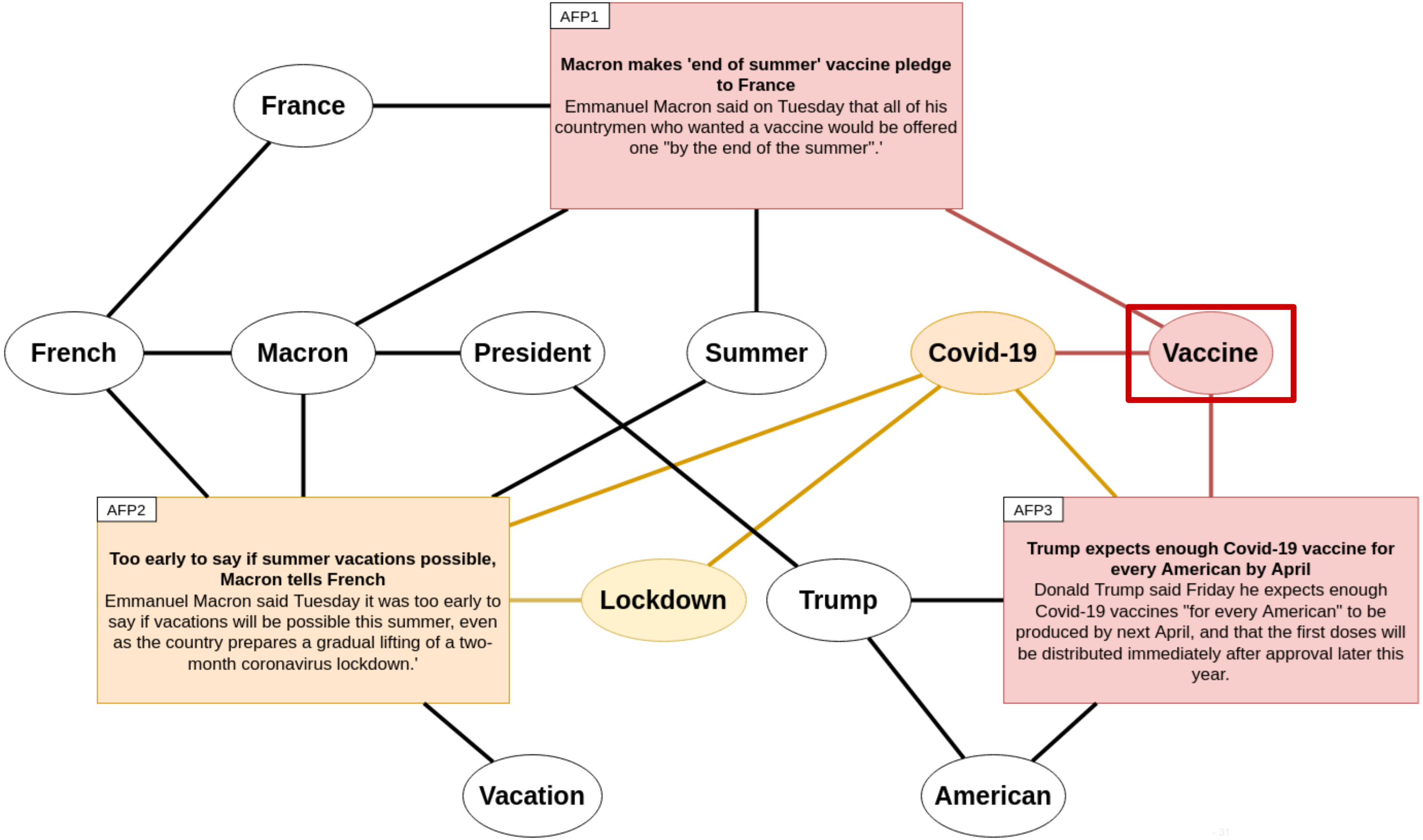
AFP2

Too early to say if summer vacations possible, Macron tells French
Emmanuel Macron said Tuesday it was too early to say if vacations will be possible this summer, even as the country prepares a gradual lifting of a two-month coronavirus lockdown.'

AFP3

Trump expects enough Covid-19 vaccine for every American by April
Donald Trump said Friday he expects enough Covid-19 vaccines "for every American" to be produced by next April, and that the first doses will be distributed immediately after approval later this year.

Text Similarity: using entities



AFP1

Macron makes 'end of summer' vaccine pledge to France
Emmanuel Macron said on Tuesday that all of his countrymen who wanted a vaccine would be offered one "by the end of the summer".'

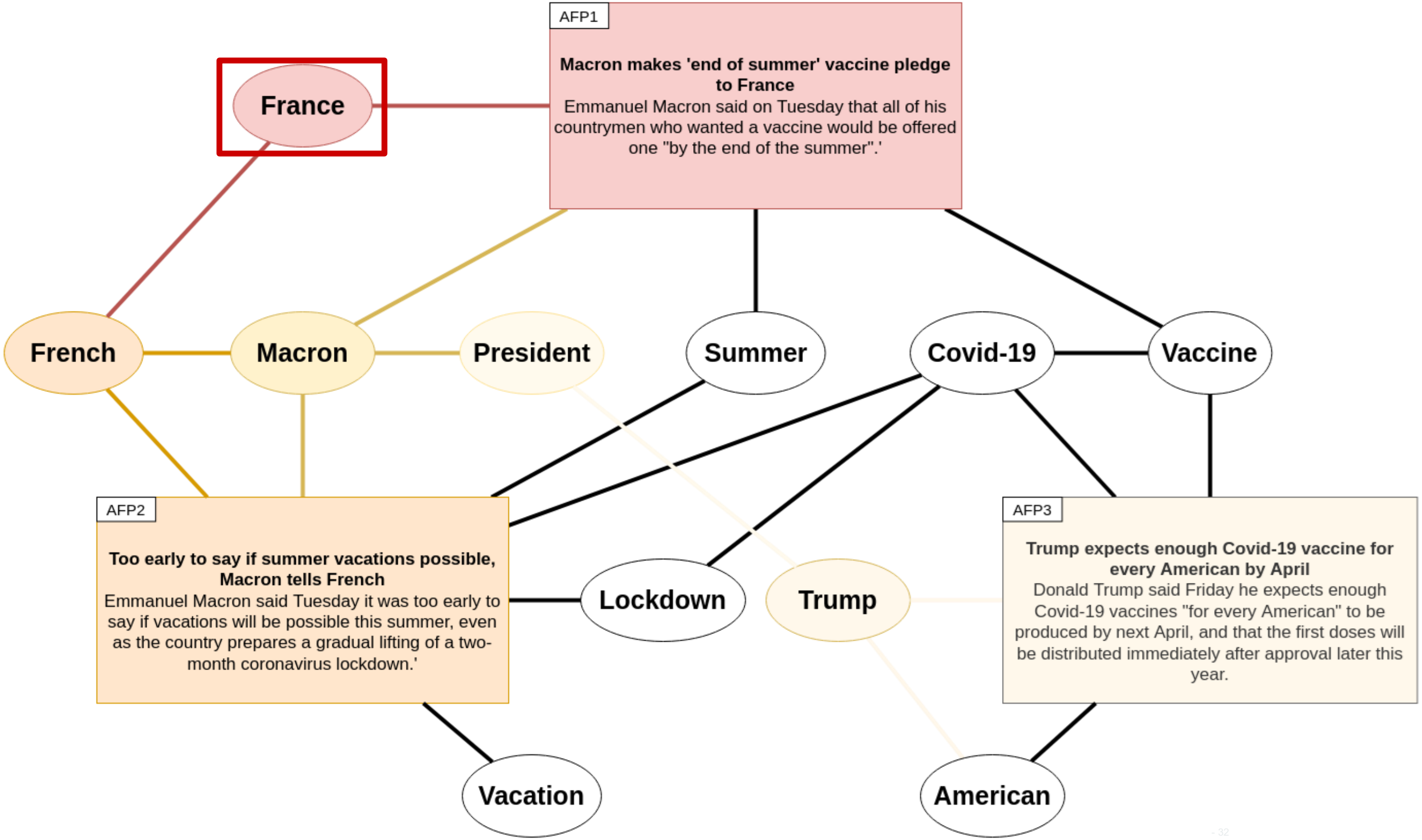
AFP2

Too early to say if summer vacations possible, Macron tells French
Emmanuel Macron said Tuesday it was too early to say if vacations will be possible this summer, even as the country prepares a gradual lifting of a two-month coronavirus lockdown.'

AFP3

Trump expects enough Covid-19 vaccine for every American by April
Donald Trump said Friday he expects enough Covid-19 vaccines "for every American" to be produced by next April, and that the first doses will be distributed immediately after approval later this year.

Text Similarity: using entities



AFP1

Macron makes 'end of summer' vaccine pledge to France
Emmanuel Macron said on Tuesday that all of his countrymen who wanted a vaccine would be offered one "by the end of the summer".'

AFP2

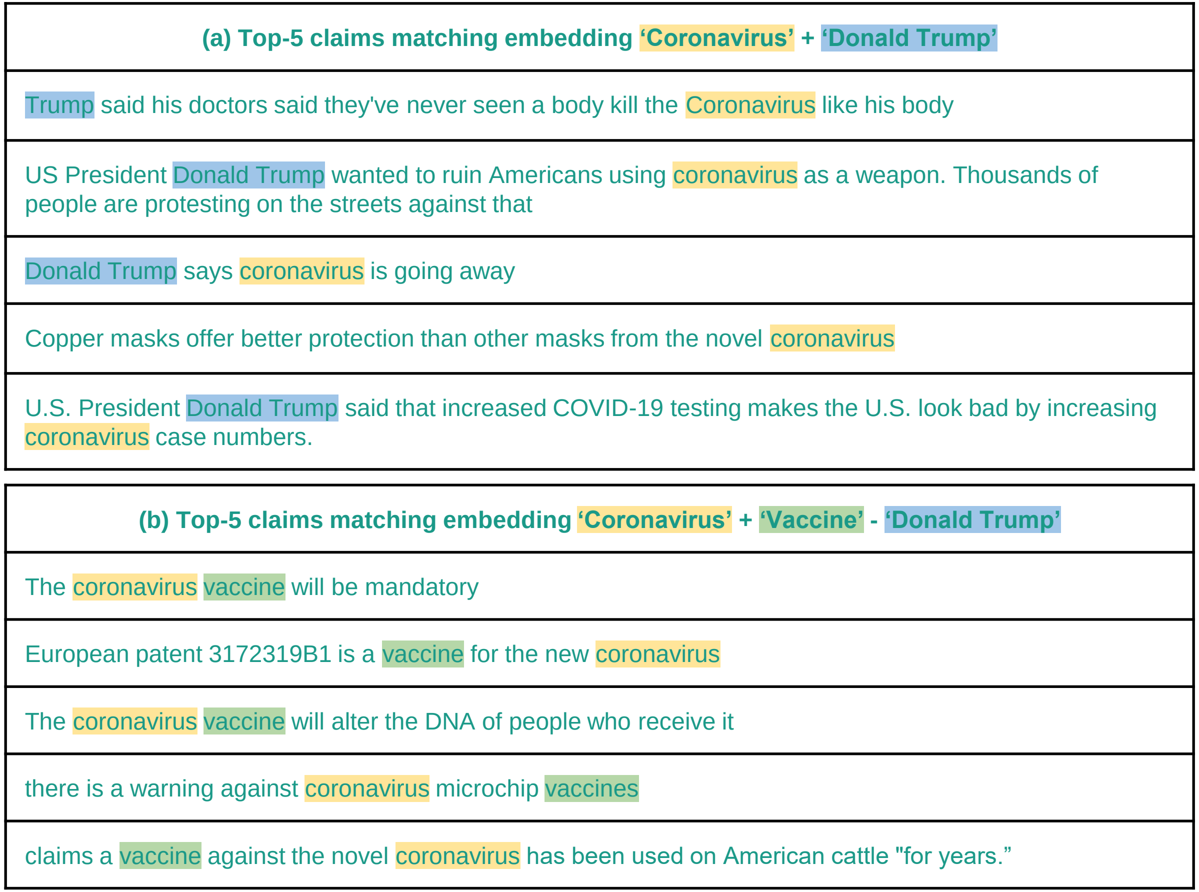
Too early to say if summer vacations possible, Macron tells French
Emmanuel Macron said Tuesday it was too early to say if vacations will be possible this summer, even as the country prepares a gradual lifting of a two-month coronavirus lockdown.'

AFP3

Trump expects enough Covid-19 vaccine for every American by April
Donald Trump said Friday he expects enough Covid-19 vaccines "for every American" to be produced by next April, and that the first doses will be distributed immediately after approval later this year.

Approach: graph-based similarity

- Node embeddings techniques (node2vec, DeepWalk, etc.)
- This allow for much flexible queries, by using arithmetic operation on embeddings



(a) Top-5 claims matching embedding ‘Coronavirus’ + ‘Donald Trump’
Trump said his doctors said they've never seen a body kill the Coronavirus like his body
US President Donald Trump wanted to ruin Americans using coronavirus as a weapon. Thousands of people are protesting on the streets against that
Donald Trump says coronavirus is going away
Copper masks offer better protection than other masks from the novel coronavirus
U.S. President Donald Trump said that increased COVID-19 testing makes the U.S. look bad by increasing coronavirus case numbers.

(b) Top-5 claims matching embedding ‘Coronavirus’ + ‘Vaccine’ - ‘Donald Trump’
The coronavirus vaccine will be mandatory
European patent 3172319B1 is a vaccine for the new coronavirus
The coronavirus vaccine will alter the DNA of people who receive it
there is a warning against coronavirus microchip vaccines
claims a vaccine against the novel coronavirus has been used on American cattle "for years.”

Explainability with Graph-based similarity

Node embeddings could also be used for factors:

‘Coronavirus’ + ‘Conspiracy Theory’ or ‘Donald Trump’ + ‘Propaganda technique’ or remove ‘Political bias’ in a tweet

They can also be used to explain the match:

AFP1 and **AFP2** match because of similar entities/concepts: **Macron**, **France**, **Summer**

AFP1 and **AFP3** match because of similar entities/concepts: **Vaccine**, **President**

Exploring Relatedness: Similarity Measures

How to improve the detection and retrieval of previously fact-checked claims?

- Semantic Textual Similarity:
  match document that share similar meaning (sentence-BERT)
- Entity-based Similarity:
  match document that share similar entities or concepts
- Factor-based Similarity:
  match documents that share similar factors
- Syntax-based Similarity:
  match documents that share similar syntax

Introducing different notions of similarity

Creation of a relatedness dataset based on fact-checks and entities:

- **Implication match**: relatedness if fact-checking information can be shared between documents
- **Narrative match**: relatedness if documents mention the same overall narrative
- **No match**
- **Implication → Narrative**

- **Entity match**: relatedness if documents mention the same entities
- **Concept match**: relatedness if documents mention the same concepts
- **No match**
- **Entity → Concept**

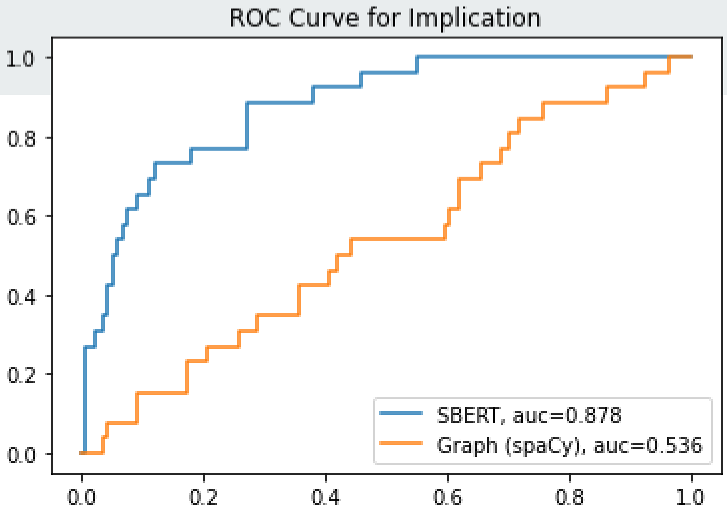
Annotation of a set of 320 pairs of tweets/claims (kappa score 0.6~0.7 on 40 pairs).
This set will serve as validation set to compare matching methods
(graph embeddings, sentence-bert, etc.)

Examples of matching

Tweet	Claim	Annotation 1	Annotation 2
Why was Tara Reade's sexual allegations against Biden dismissed and not investigated	CNN hasn’t published a single article about former Senate staffer Tara Reade’s sexual assault allegation against former Vice President Joe Biden	Implication	Entity
@Debgolf2 @wileynickel @NCSenateDems 80,000 didn’t get guns via the permitting system. It works	We "have laws on the books designed to prevent people with mental illnesses from getting firearms."	Narrative	Concept
why are vaccines being pushed, if there are proven effective cures available on the market for 10 to 60 years? #hyroxychloroquid #vermectin etc. #Covid19VaccineReport	COVID-19 no more lethal than flu there is no pandemic COVID vaccines are unsafe and ineffective and 2020 death rate no higher than average.	Narrative	None
Illegal border crossings are out of control thanks to President Biden's border policies. The #BidenBorderCrisis is real, and our 'border czar' @VP is nowhere in sight.	the image show migrants in a holding facility in 2021 during President Joe Biden"s tenure	Narrative	Concept
@BarstoolPU You should know better than to share propaganda like this. Let people make their own decisions! There is plenty of evidence that the vaccine causes diseases such as AIDS and Down syndrome.	A "COVID-19 Vaccine Q&A" includes a series of claims, including that the vaccines currently in use skipped animal testing, contain cells from foreign sources such as monkeys and cause serious harm to human health.	Implication	Concept
Do not be surprised if we learn in the days ahead that the Trump rioters were infiltrated by leftist extremists. Note: this is not to excuse any of them	"In office, President Trump has accomplished more in his first 100 days than any other President since Franklin Roosevelt	None	Concept

Predicting « implication » match

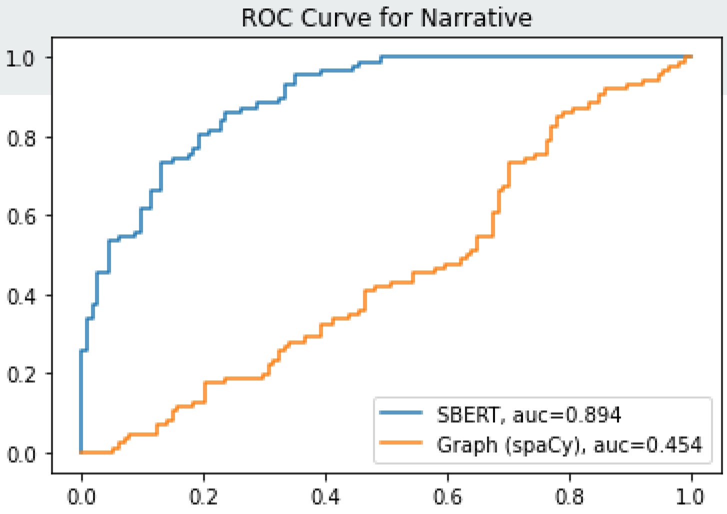
26 implication matches



Method	nodes	edges	MRR	Acc@1	Acc@5	Acc@10	Acc@50	Acc@100
Sentence-BERT	0	0	**0.844**	**0.808**	**0.885**	**0.923**	**0.962**	**1**
spaCy	20,814	56,223	0.44	0.346	0.538	0.615	0.692	0.808
spaCy + factors	20,849	184,135	0.271	0.192	0.385	0.423	0.654	0.731

Predicting « narrative » match

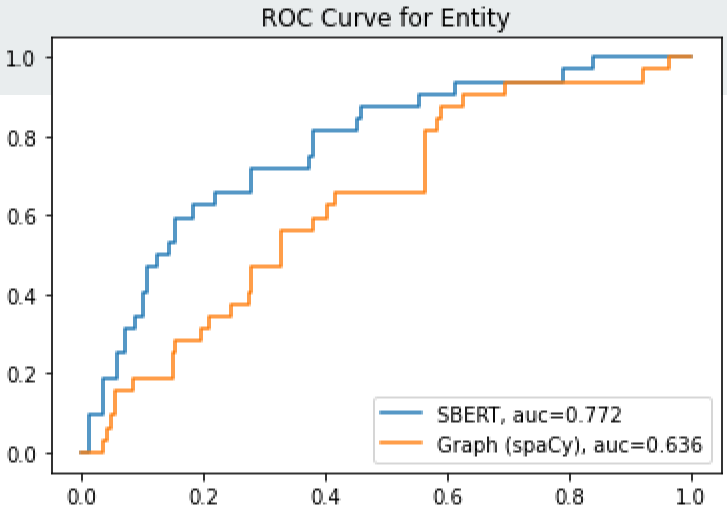
86 narrative matches



Method	nodes	edges	MRR	Acc@1	Acc@5	Acc@10	Acc@50	Acc@100
Sentence-BERT	0	0	**0.783**	**0.744**	**0.814**	**0.884**	**0.953**	**0.977**
spaCy	20,814	56,223	0.401	0.349	0.442	0.488	0.628	0.721
spaCy + factors	20,849	184,135	0.184	0.128	0.244	0.314	0.523	0.581

Predicting « entity » match

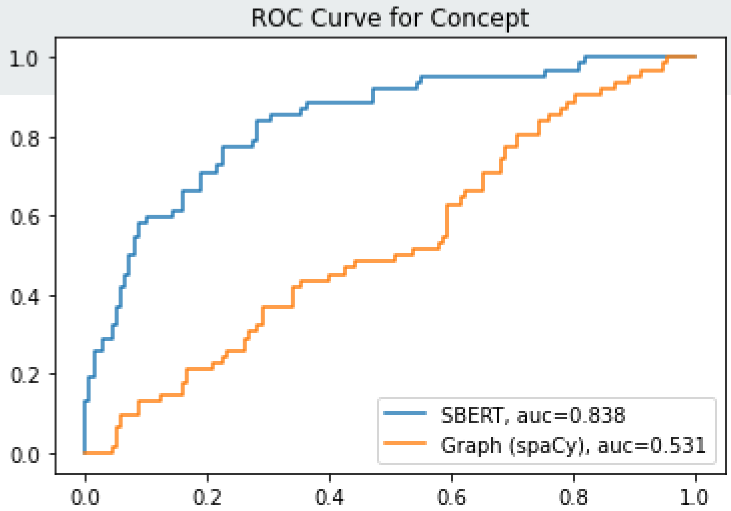
32 entity matches



Method	nodes	edges	MRR	Acc@1	Acc@5	Acc@10	Acc@50	Acc@100
Sentence-BERT	0	0	**0.732**	**0.688**	**0.75**	**0.781**	**0.938**	**0.969**
spaCy	20,814	56,223	0.578	0.469	0.656	0.781	0.844	0.906
spaCy + factors	20,849	184,135	0.283	0.188	0.406	0.5	0.781	0.812

Predicting « concept » match

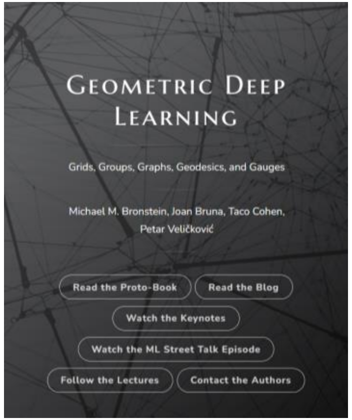
62 concept matches



Method	nodes	edges	MRR	Acc@1	Acc@5	Acc@10	Acc@50	Acc@100
Sentence-BERT	0	0	**0.765**	**0.726**	**0.79**	**0.839**	**0.935**	**0.968**
spaCy	20,814	56,223	0.465	0.403	0.5	0.565	0.645	0.758
spaCy + factors	20,849	184,135	0.247	0.194	0.306	0.371	0.532	0.597

Take Away

- **Transformer-based models to predict factors related to misinformation work “reasonably” when having enough training data!**
  - ⚠ For generalization across document genres and domains
- **Large Language Models are cheap labelers!**
  We propose to compress few-shot learning into explicit definitions
- **Knowledge Graphs as data structures to anchor factuality and to explain predictions**
  - node2vec/RDF2Vec/GNN: how far Geometric Deep Learning brings us?
  - ULTRA: A Foundation Model for KG Reasoning (github)
- **What do we need going forward?**
  - Structured reasoning and knowledge capture
    but with more language-like representations



Geometric Deep Learning

Grids, Groups, Graphs, Geodesics, and Gauges

Michael M. Bronstein, Joan Bruna, Taco Cohen, Petar Veličković

@rtroncy
raphael.troncy@eurecom.fr

EURECOM

Thank you!

Code: /D2KLab/