

# Explainable Fact Checking for Statistical and Property Claims

Paolo Papotti



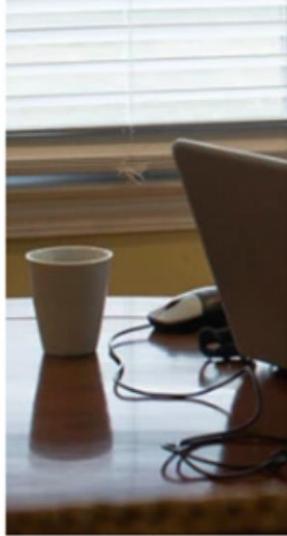
RuleML General Telecon- 28<sup>th</sup> April 2021



INDEP

Voices

# Fake news references violence. What the



Cameron Harris is an electrical worker. Demczuk for The New York Times



Rohingya refugees in Bangladesh, in 2017, keep its platform from Myanmar. Adam Dean

By Alexandra Stevens

Nov. 6, 2018

[https://www.nytimes.com/2017/01/18/us/fake-news-hillary-clinton-cameron-harris.html?\\_r=0](https://www.nytimes.com/2017/01/18/us/fake-news-hillary-clinton-cameron-harris.html?_r=0) gove-boris-johnson-theresa-may-a7533806

## Global Pediatric Health



*Glob Pediatr Health.* 2019; 6: 2333794X19862949.

PMCID: PMC6657116

Published online 2019 Jul 24.

PMID: 31384629

doi: [10.1177/2333794X19862949](https://doi.org/10.1177/2333794X19862949)

### Anti-Vaccine Decision-Making and Measles Resurgence in the United States

Olivia

Measles containing vaccine coverage second dose\*, 2016

Au

0–84%

This

85–94%

Intro

95–99%

Meas

No data

to ot

Not included

and J

psych

State

resea

healt

for th

Epic

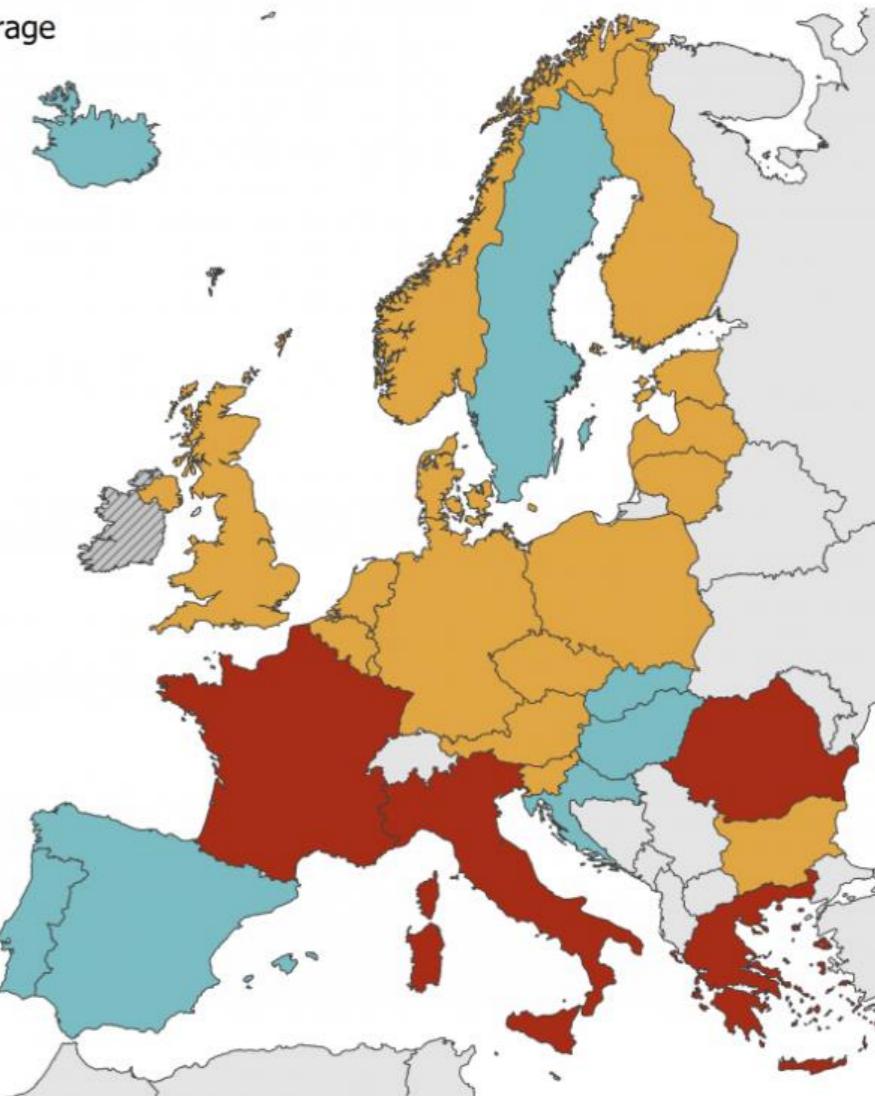
Luxembourg

Duri

Malta

State

Malta



\*WHO, reported estimates

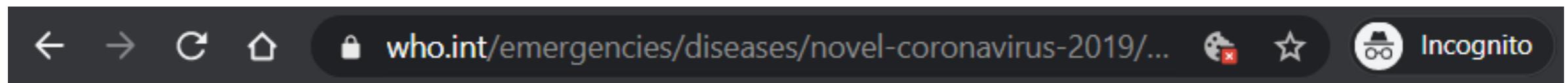
ECDC

# COVID-19 Infodemic

“The outbreak has been accompanied by a massive *infodemic* - an over-abundance of information – some accurate and some not – that makes it hard for people to find trustworthy sources and reliable guidance when they need it.”



**World Health Organization**



To help alleviate suffering and save lives, WHO has been working night and day in five key ways:

- 1. Helping build countries' capacity to prepare and respond
- 2. Providing accurate information and fight the infodemic, together with numerous partners
- 3. Ensuring supplies of essential medical equipment for frontline health workers.
- 4. Training and mobilizing health workers.
- 5. Accelerating research and development.

<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>

# Checking claims



sun kills coronavirus



CORONAVIRUS WHICH TAKES OFF FADEST REMAINS FOR 9 HOURS SO WASHING CLOTHES ...

sites.nationalacademies.org › covid-sunscreen ▾

## Does sunlight kill the coronavirus? | National Academies

Claim: Since sunlight can kill the coronavirus, wearing sunblock is a bad idea during the coronavirus pandemic.

Fact check by National Academies: False. Scientists are still studying...

Feedback

Jeff Smith shared a link

Additional Reporting On This

Before you share this content, you might want to know that there is additional reporting on this from Snopes.com, Factcheck.org and Associated Press.

Visit Snopes.com

Visit Factcheck.org

Visit Associated Press

[Cancel](#) [Post Anyway](#)

Donald J. Trump @realDonaldTrump · 9h  
There is NO WAY (ZERO!) that Mail-In Ballots will be anything less than substantially fraudulent. Mail boxes will be robbed, ballots will be forged & even illegally printed out & fraudulently signed. The Governor of California is sending Ballots to millions of people, anyone.....

Get the facts about mail-in ballots

22K 24.3K 86.7K

Donald J. Trump @realDonaldTrump · 9h  
....living in the state, no matter who they are or how they got there, will get one. That will be followed up with professionals telling all of these people, many of whom have never even thought of voting before, how, and for whom, to vote. This will be a Rigged Election. No way!

Get the facts about mail-in ballots

# Fact checking

The screenshot shows a map of North America with specific locations marked: Winnipeg, Saint Paul (with a pin labeled '2'), United States (with a red checkmark pin), Oklahoma City, Austin, and Bahamas. A sidebar menu on the right includes options for 'LATEST NEWS', 'FACT-CHECKING' (which is currently selected), 'TECH & CHECK', and 'ABOUT THE LAB'. A callout box highlights 'Active (304)' with a checked checkbox, and 'Inactive (100)' with an unchecked checkbox.

**Problem:**  
Manual fact checking cannot  
scale to the amount of  
information produced everyday

<https://www.poynter.org/coronavirusfactsalliance/>

# THE WALL STREET JOURNAL.

English Edition ▾ | May 25, 2020 | Print Edition | Video

Home World U.S. Politics Economy Business **Tech** Markets Opinion Life & Arts Real Estate WSJ. Magazine

VIRUS

FINDING A JOB BACK TO WORK Q&A GOING OUTSIDE SAFELY CAN I GET SICK AGAIN? CHILDCARE & SUMMER CAMPS CAN I TRAVEL? STATE REOPENINGS  
TESTS & TREATMENTS WHAT 6 FEET LOOKS LIKE EN ESPAÑOL

OTICE

We use cookies for analytics, advertising and to improve our site. You agree to our use of cookies by closing this message box or continuing to use our site. To including how to change your settings, see our [Cookie Notice](#)

TECH

## Facebook's Fact Checkers Fight Surge in Fake Coronavirus Claims

Contractors battle bogus assertions about canine vaccines and free baby formula; 'We've maxxed out'



RECOMMENDED VIDE

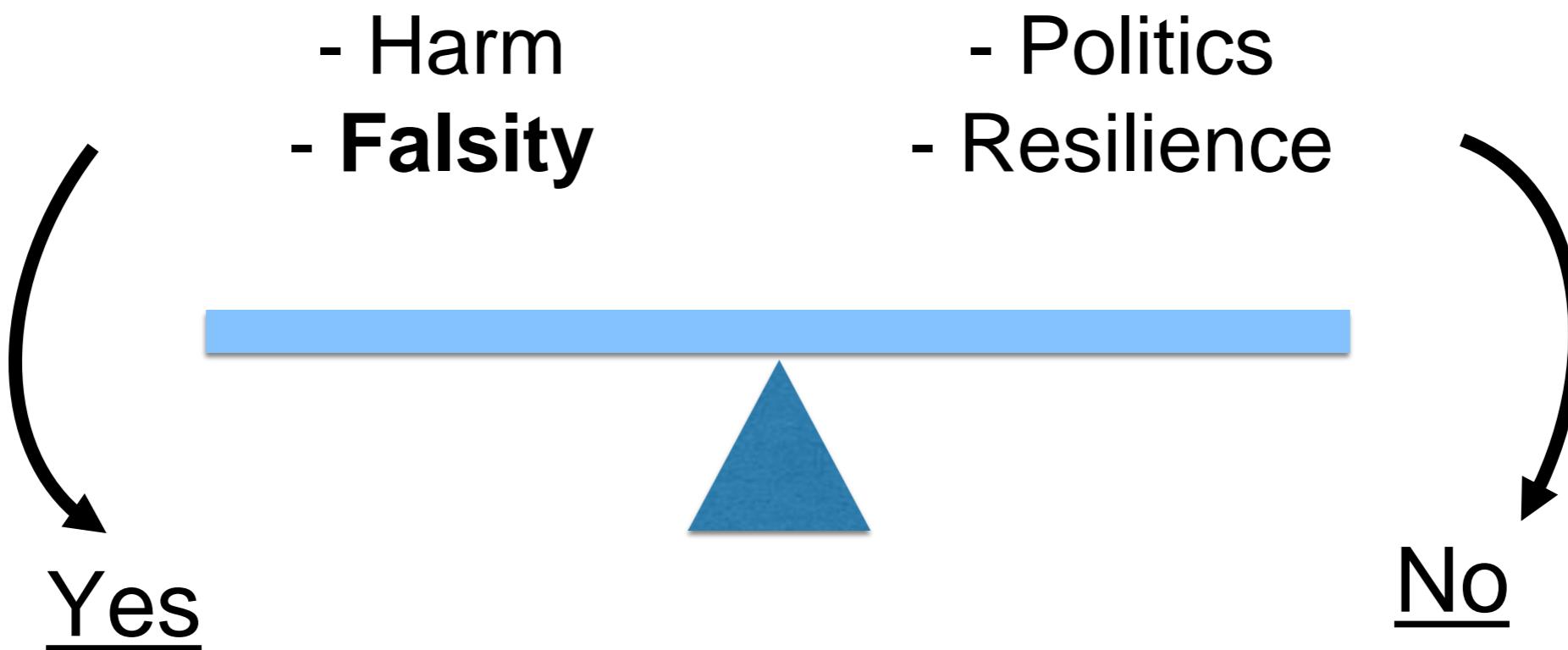
1. Coronavirus Updat China Scraps GDP Target, Trump's Swine-State Tour

<https://www.wsj.com/articles/facebook-fact-checkers-fight-surge-in-fake-coronavirus-claims-11585580400>



# Disclaimer: no computational solution for content moderation

- Intervention?



<https://regulate.tech/partisan-the-missing-word-26th-april-2020/>

# Computational fact checking

- The general problem:  
Given some content, assess if it is true or false

# SIGIR'19 Workshop: ROME 2019



Digital and Technology

Health & Social Care

Local Government & Communities

Police & Criminal Justice

The Tackling Misinformation Conference 2021: Communicating Effectively to Improve Local Services

Date: July 7, 2021

Location: Online

Event Type: Online



THE WEB  
CONFERENCE

## Workshop on Misinformation Integrity in Social Networks

**EU Fact Checking & EU Fact Finding**  
The EU at your fingertips!

7-8 October 2021  
Brussels

28-29 October 2021  
Online

## Call for Papers

## 2021 Workshop on NLP4IF:

**FOR TRUTH AND TRUST**

**ROMCIR 2021**  
Workshop on REDUCING ONLINE MISINFORMATION  
THROUGH CREDIBLE INFORMATION RETRIEVAL – Lucca |

Tuscany | Italy (Online Event) | April 1, 2021

EDMO Workshop – Artificial Intelligence and fact-checking:  
promises and perils

**SIGKDD 2019 Workshop, August 5th, Anchorage, Alaska**



**ABOUT MISINFOCON**



# Computational fact checking

- The general problem:  
Given some textual content, assess if it is true or false
- Our problem:  
Given some **textual content** and some reference **data**,  
assess if the content is **true or false** and explain **why**
  - **# identifying check-worthy claims** [ClaimBuster]
  - **# matching claims to existing checks** [FullFact, ClaimReview]
  - **# model trust of the sources** [Tanbih]
- Effective, Scalable, Interpretable

# Fact checking

A screenshot of a Google search results page. The search query is "Yassine Belattar pancarte "Nous vous coloniserons"". The results show approximately 306 results found in 0.31 seconds. The top result is a link to Liberation.fr about a protest against Islamophobia. Below it is a result from AFP Factuel stating that Yassine Belattar did not hold a sign reading "Nous vous coloniserons". A blue oval highlights the "Fact check by AFP Factuel: Faux" link. To the right, a modal window titled "Additional Reporting On This" provides links to Snopes.com, Factcheck.org, and Associated Press for further reporting.

Yassine Belattar pancarte "Nous vous coloniserons"

All Images News Videos Maps More Settings Tools

About 306 results (0.31 seconds)

[Manif contre l'islamophobie : attention à la manipulation](#)  
https://www.liberation.fr › checknews › 2019/1  
6 days ago - «Avec vos lois démocratiques, nous vous colonisez», déclara lundi 14 mai l'humoriste Yassine Belattar, et l'auteur de la phrase a été viré de Twitter. L'humoriste Yassine Belattar était effectivement présent à la manifestation contre l'islamophobie organisée à Paris le 14 mai.

[Non, l'humoriste Yassine Belattar n'a pas brandi une pancarte avec la phrase "Nous vous coloniserons"](#)  
https://factuel.afp.com › non-lhumoriste-yassine-belattar-n-a-pas-brandi-une-pancarte-avec-la-phrase-nous-vous-coloniserons

Claim: L'humoriste Yassine Belattar brandit une pancarte avec la phrase "Nous vous coloniserons".  
Claimed by: Sources multiples  
Fact check by AFP Factuel: Faux

Feedback

**Additional Reporting On This**

Before you share this content, you might want to know that there is additional reporting on this from Snopes.com, Factcheck.org and Associated Press.

[Visit Snopes.com](#)

[Visit Factcheck.org](#)

[Visit Associated Press](#)

Cancel Post Anyway

# Computational fact checking

- Given some **textual content** and some reference **data**, assess if the content is **true or false** and explain **why**
  - `Check(claim,data) =`  
true/false label,  
confidence value,  
description of the subset of the data that implies the decision
  - **Property** and statistical claims

# Checking property claims

- Input textual claim:  
“Elon Musk is the founder of Chevrolet”
- Output: FALSE because  
“Chevrolet was *founded* in 1911 and Elon Musk was *born* in 1971”

# Checking property claims

Not in data. Count occurrences/text patterns/... on Web

“Chevrolet”

- Output: FALSE because

“Chevrolet was *founded* in 1911 and Elon Musk was *born* in 1971”

Concepts of person, company, birth year, founding year, logical contradiction

# Checking property claims

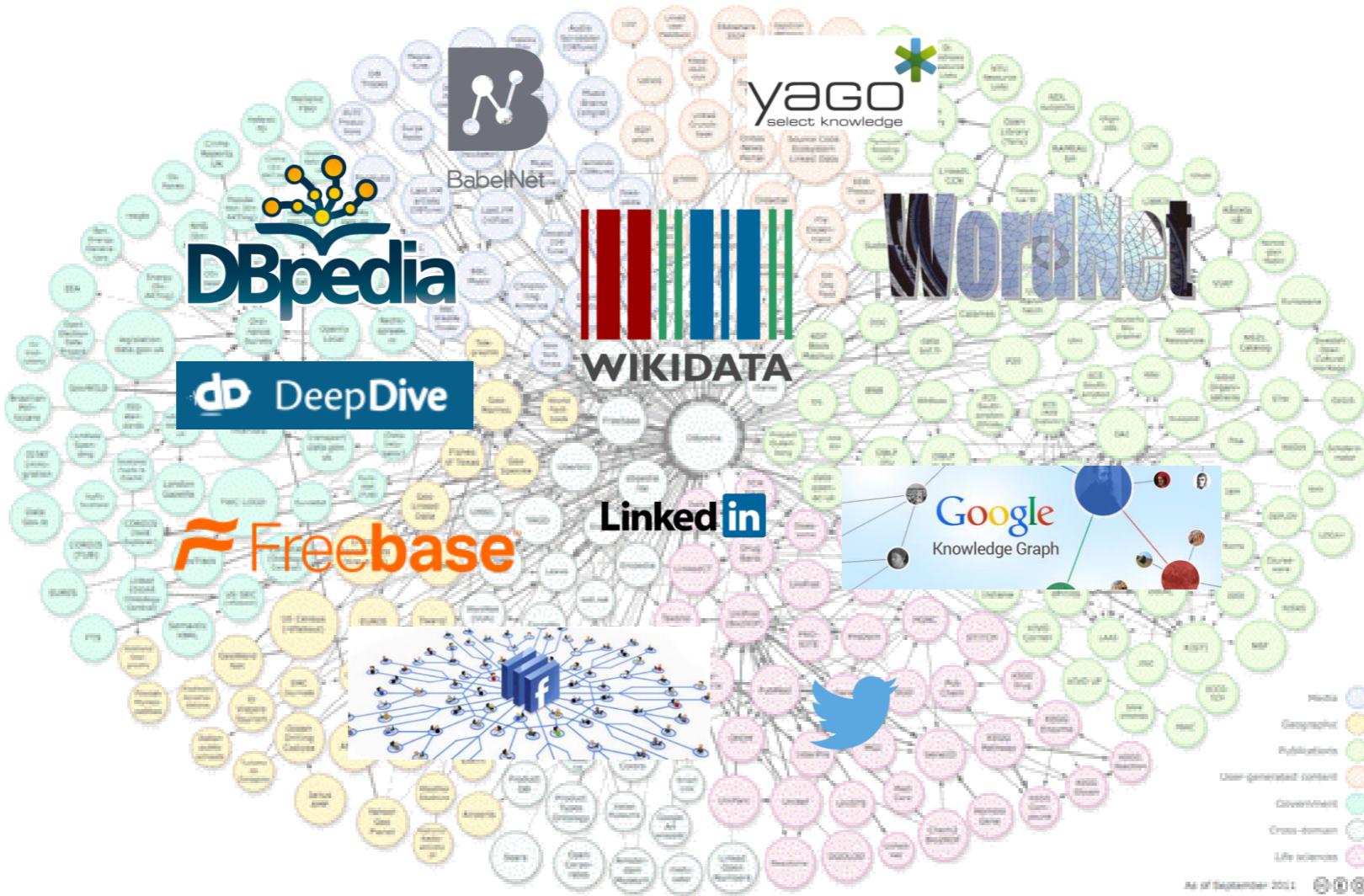
- Input textual claim:  
“Michael White’s alma mater is UT Austin”
- Output: FALSE because  
“M. White *works* at UT Austin and  
has Abilene Christian University and Yale  
Divinity School as *alma mater*”

# Checking property claims

- Input
  - Textual claim: “Michael White’s alma mater is UT Austin” →  
*almaMater(Michael White, UT Austin)*

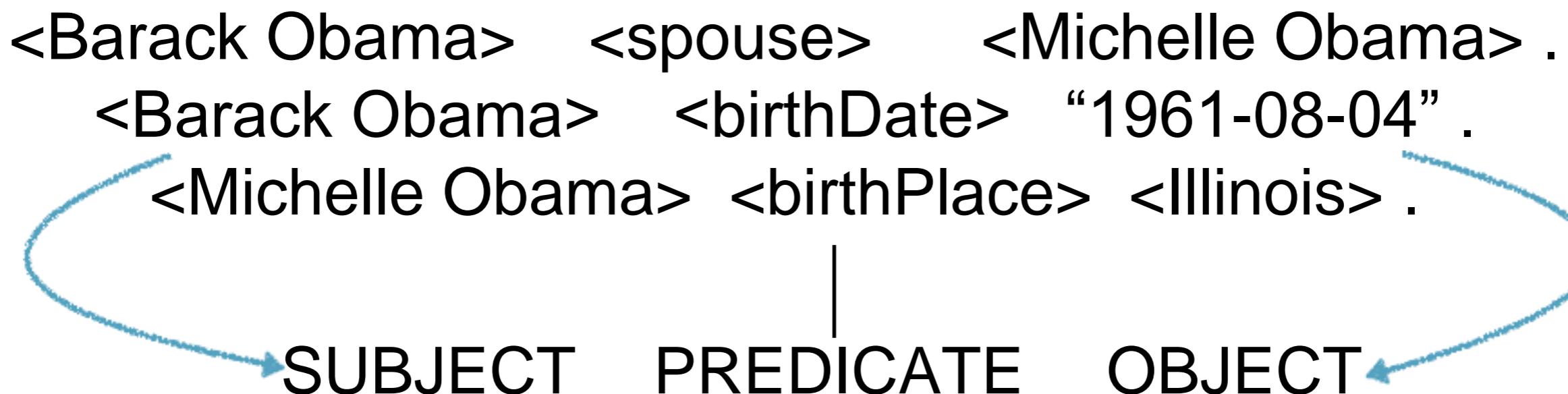


# RDF Knowledge Graphs



Walmart, Amazon, SAP, Amadeus, ...

# RDF KGs



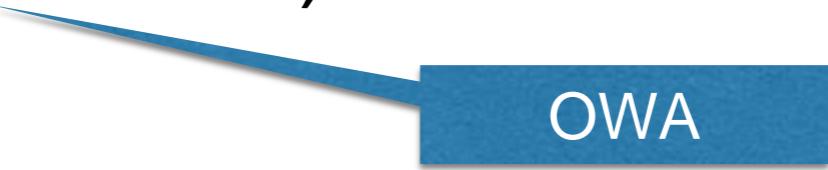
Name	# Entity types	# Entity instances	# Relation types
Knowledge Vault (KV)	1100	45M	4469
DeepDive [21]	4	2.7M	34
NELL [6]	271	5.19M	306
PROSPERA [20]	11	N/A	14
YAGO2 [16]	350,000	9.8M	100
Freebase [4]	1,500	40M	35,000
Knowledge Graph (KG)	1,500	570M	35,000

[Dong and Srivastava, 2015]

# Computational fact checking

- Given some **textual content** and some reference **data**, assess if the content is **true or false** and explain **why**

- Check(claim, *incomplete KG*) =  
true/false label,  
confidence value,  
description of the subset of the data that implies the decision



OWA

→ how to assess/explain if claim is **NOT** in KG?

# Explaining the decision

- Output
  - “Michael White’s alma mater is UT Austin” is FALSE because
  - Evi 1: employer(Michael White, UT Austin)  
Evi 2: almaMater(Michael White, Abilene Christian Univ.)  
almaMater(Michael White, Yale Divinity School)
  - Approximate rules over the KG
    - 0.1: almaMater(x,y), almaMater(x,z),  $x \neq z \Rightarrow \perp$
    - 0.3: almaMater(x,y), employer(x,y)  $\Rightarrow \perp$

# Negative Rules

$$0.9 \text{ foundingYear}(x,z) \wedge \text{birthYear}(y,w) \wedge (z \leq w) \wedge \text{founder}(x,y) \Rightarrow \perp$$

Atom = predicate from KG

[*claim*] *founder(Chev., E.Musk)*  
[KG] *foundingYear(Chev, 1911)*  
[KG] *birthYear(E.Musk, 1971)*

$x=\text{Chev.}, z=1911, \dots$   
*At least one triple is incorrect*

“negative” rules identify inconsistencies

$$1.0 \text{ President}(x, \text{USA}) \wedge \text{PlaceOfBirth}(x,y) \wedge y \neq \text{USA} \Rightarrow \perp$$
$$0.9 \text{ isMarriedTo}(x,y) \wedge \text{hasChild}(x,y) \Rightarrow \perp$$

# Horn Rules Discovery

0.7  $\text{hasChild}(x,z) \wedge \text{hasChild}(y,z) \Rightarrow \text{spouse}(x,y)$

*child(Barack, Sasha) \wedge child(Michelle, Sasha)*  
 $\Rightarrow \text{spouse}(Barack, Michelle)$

**“positive” rules** address incomplete data

1.0  $\text{notableWork}(y,x) \Rightarrow \text{creator}(x,y)$  (Wikidata)  
0.8  $\text{hasChild}(z,y) \wedge \text{isMarriedTo}(x,z) \Rightarrow \text{hasChild}(x,y)$  (Yago)

[Galárraga et al. WWW 2013]

# ExpClaim system

“Elon Musk is the founder of Chevrolet” →  
founder(Chevrolet, E.Musk)

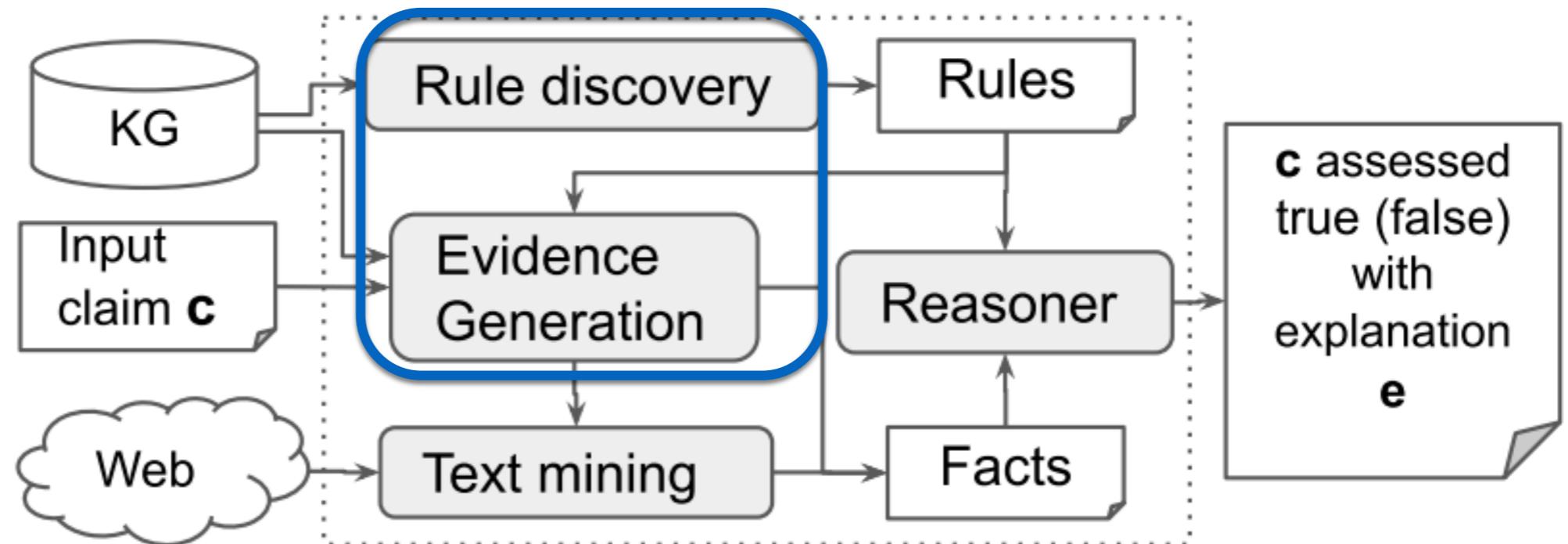


Ahmadi et al. *Explainable Fact Checking with Probabilistic Answer Set Programming*. TTO 2019.  
<https://github.com/ppapotti/expclaim>



# Mine Logical Rules

founder(Chevrolet, E.Musk)



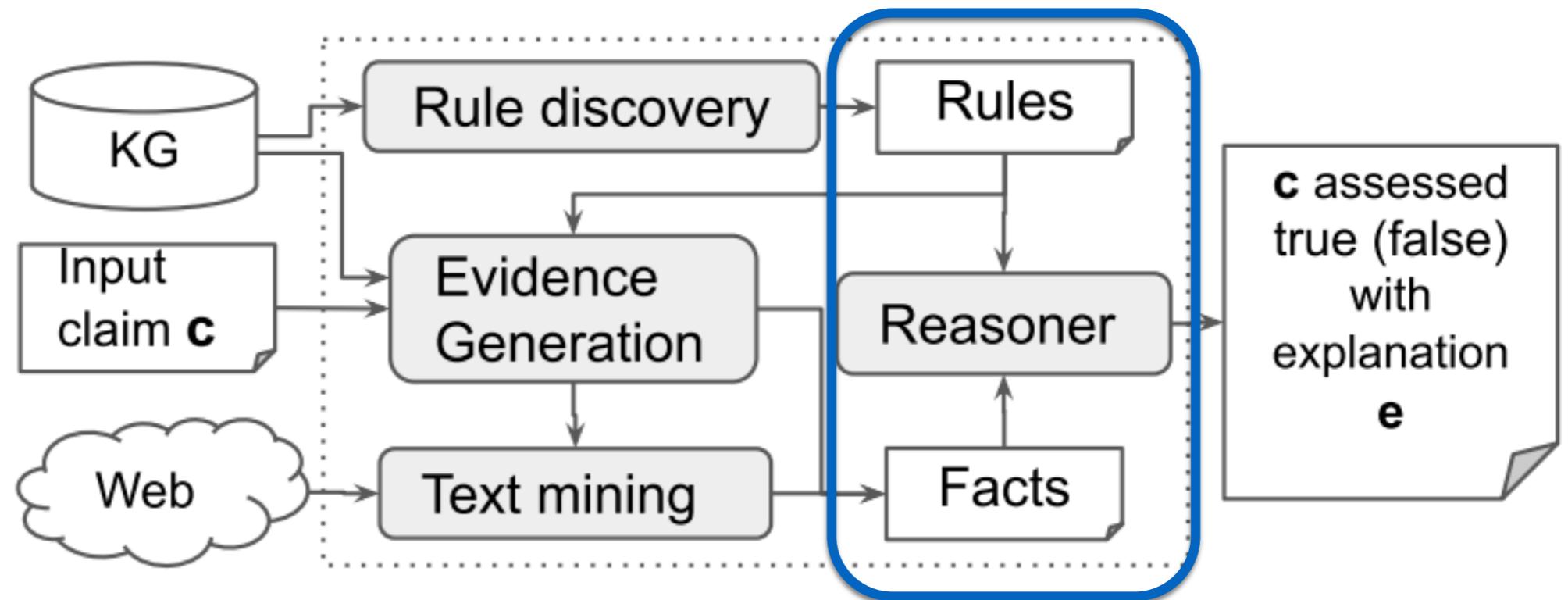
Gather rules and evidence (facts) from the KG

$$0.9 \text{ foundingYear}(x,z) \wedge \text{birthYear}(y,w) \wedge (z \leq w) \wedge \text{founder}(x,y) \Rightarrow \perp$$

*[K] foundingYear(Chev, 1911), [K] birthYear(E.Musk, 1971), [c] founder(E.Musk, Chev)*

# MAP Reasoning

“Elon Musk is the founder of Chevrolet”



A probabilistic extension of answer set programs with the concept of **weighted rules** derived from Markov Logic

[Lee et al., KR 2016]

- ```

FALSE : almaMater(Michael White, UT Austin)
         ← employer(Michael White, UT Austin)
         ← occupation(Michael White, UT Austin)
         ← almaMater(Michael White, Abilene
                      Christian Univ.), almaMater(Michael
                      White, Yale Divinity School)
      
```
- Baselines: text miner (CredE), link prediction (KGM), hard rules (ASP)

|        | <i>almaMat.</i> | <i>deathPl.</i> | <i>spouse</i> | <i>vicePres.</i> |
|--------|-----------------|-----------------|---------------|------------------|
| CRED E | .41(.03)        | .59(.06)        | .44(.07)      | .36(.15)         |
| KGM    | .73(.08)        | .68(.01)        | .86(.01)      | .81(.03)         |
| ASP    | .70(.06)        | .01(.01)        | .31(.08)      | .18(.16)         |
| MAP    | .88(.14)        | .75(.15)        | .87(.11)      | .66(.22)         |
| MAP+W  | .88(.09)        | .83(.11)        | .86(.10)      | .68(.18)         |

# Computational fact checking

- Given some **textual content** and some reference **data**, assess if the content is **true or false** and explain **why**
  - `Check(claim,data) =`  
true/false label,  
confidence value,  
description of the subset of the data that implies the decision
  - Property and **statistical** claims

# Checking statistical claims

Liz Wheeler ✅ @Liz\_Wheeler · Mar 20

U.S. death rate isn't 3.4%. It's 1.3% & falling.

Social distancing won't stop spread of COVID-19 unless  
(Won't happen.)

Economic damage is ALREADY hurting people.

So, in 2 months, we'll see economic devastation AND re-  
does that make sense?

386

1.4K

4.4K

- Output: FALSE

<https://coronacheck.eurecom.fr> /en, /i

How did we verify your statement? X

**Claim:** 0.013000000000000001

**Inferred formula:** a/b

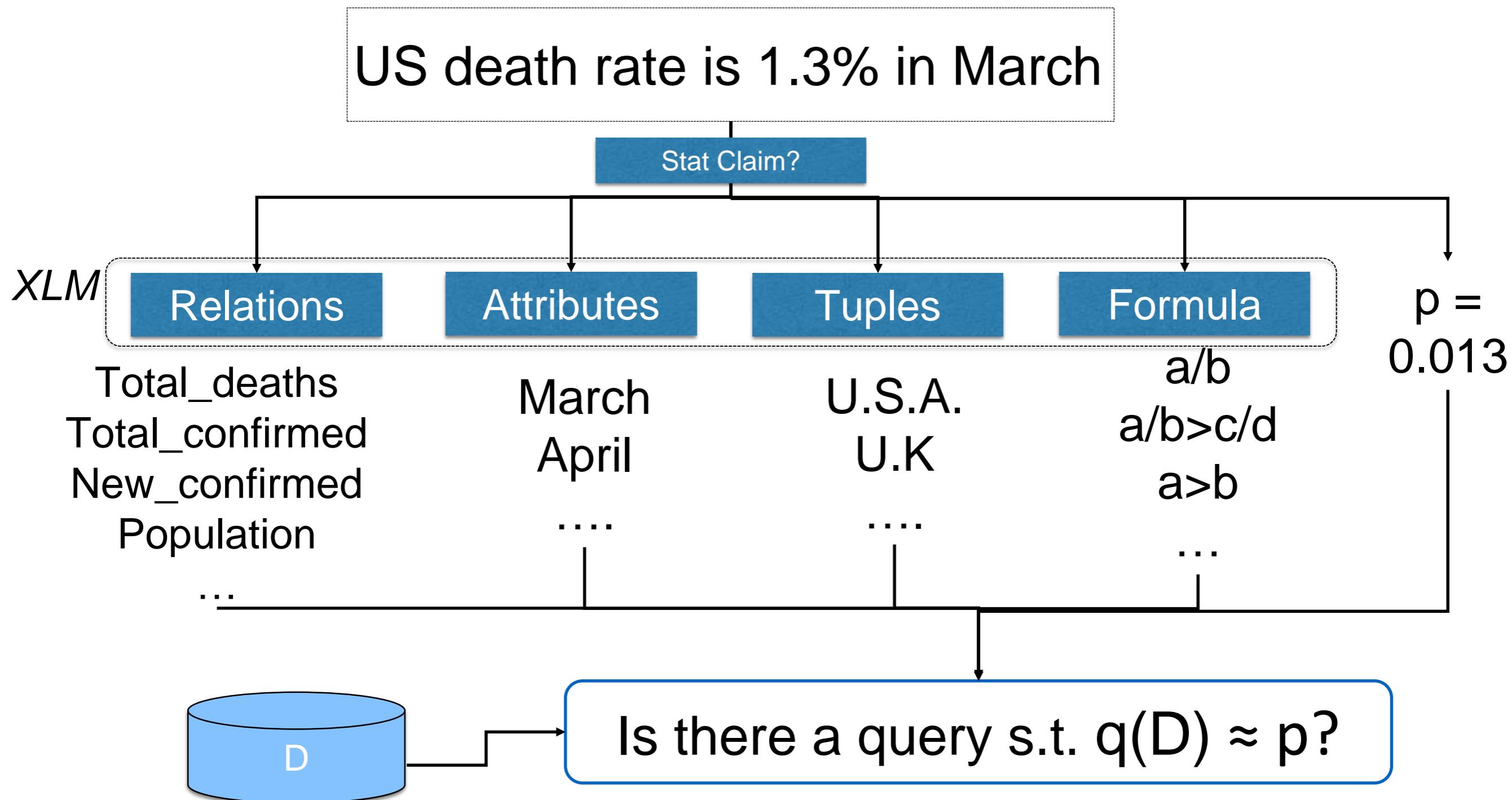
**Inferred formula value:** 0.028521777946319233

**List of variables:**

| Variable name | a            |
|---------------|--------------|
| Table         | total_deaths |
| Month         | March        |
| Country       | USA          |
| Value         | 5367         |

| Variable name | b               |
|---------------|-----------------|
| Table         | total_confirmed |
| Month         | March           |
| Country       | USA             |
| Value         | 188172          |

# From text to query elements to data driven check



# User feedback

- “Yesterday the number of new cases in France decreased **a lot**”

New\_confirmed,  
14-15 Sept,  
France,  
 $(b-a)/a < ?$

- “In **Europe**, the number of new cases is decreasing this month”

New\_confirmed,  
Aug-Sept,  
*tuple* ?  
 $(b-a)/a < 0$

We are unable to verify your statement, please help us by picking the best matches out of the following options:

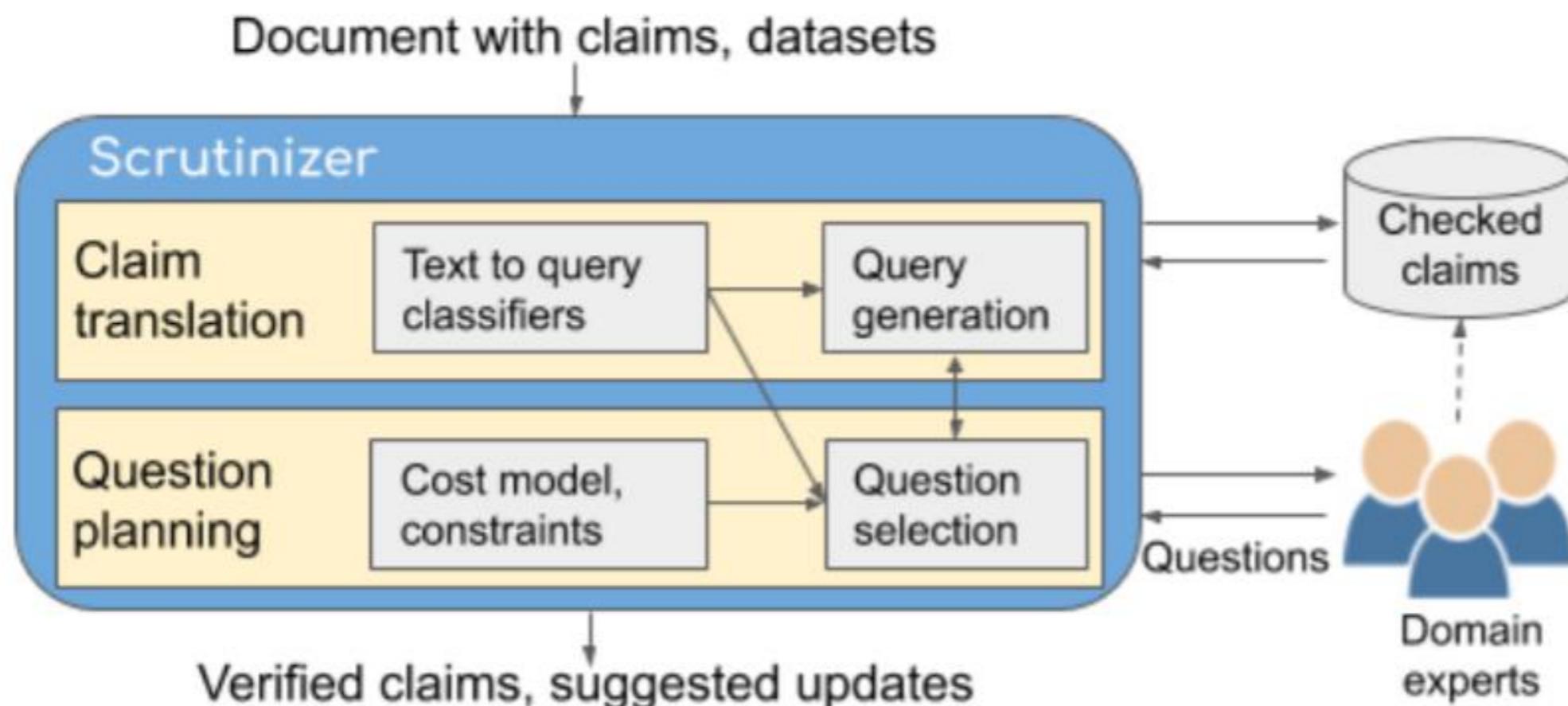
- total\_deaths-total\_confirmed
- new\_confirmed-total\_confirmed
- total\_deaths-total\_recovered
- total\_recovered-total\_deaths
- total\_recovered-total\_confirmed

Can you suggest any other dataset?

Close

Next

# Scrutinizer



# Ready for adoption?

[IJCAI'21 survey <https://arxiv.org/pdf/2103.07769.pdf> ]

## Automated Fact-Checking for Assisting Human Fact-Checkers

**Preslav Nakov<sup>1\*</sup>, David Corney<sup>2</sup>, Maram Hasanain<sup>3</sup>, Firoj Alam<sup>1</sup>, Tamer Elsayed<sup>3</sup>, Alberto Barrón-Cedeño<sup>4</sup>, Paolo Papotti<sup>5</sup>, Shaden Shaar<sup>1</sup>, Giovanni Da San Martino<sup>6</sup>**

<sup>1</sup>Qatar Computing Research Institute, HBKU, Qatar, <sup>2</sup>Full Fact, UK, <sup>3</sup>Qatar University, Qatar,

<sup>4</sup>Università di Bologna, Italy, <sup>5</sup>EURECOM, France, <sup>6</sup>University of Padova, Italy

{pnakov, fialam, sshaar}@hbku.edu.qa, david.corney@fullfact.org, a.barron@unibo.it,  
{maram.hasanain, telsayed}@qu.edu.qa, papotti@eurecom.fr, dasan@math.unipd.it

### Abstract

The reporting and analysis of current events around the globe has expanded from professional, editor-lead journalism all the way to citizen journalism. Politicians and other key players enjoy direct access to their audiences through social media, bypassing the filters of official cables or traditional media. However, the multiple advantages of free speech and direct communication are dimmed by the misuse of the media to spread inaccurate or misleading claims. These phenomena have led to the modern incarnation of the *fact-checker* — a professional whose main aim is to examine claims using

Such organizations are also potential beneficiaries of or leaders in automated fact-checking research. As misinformation became a major concern globally, tech companies, national and international agencies began work in the area. Recently, several international initiatives have also emerged such as the *Credibility Coalition*<sup>5</sup> and *EUfactcheck*<sup>6</sup>. Along side some tools have also been made available such as Google Factcheck<sup>7</sup> and Hoaxy<sup>8</sup>. Moreover, fact-checking is a common task in settings that go beyond online misinformation, as the verification of content accuracy is a priority for many organizations [Karagiannis *et al.*, 2020].

A large body of research is devoted to developing automatic systems for fact-checking [Vo and Lee, 2018; Shu



Yesterd

A few mins ago I was  
 "information you have  
 clicked the pop-up  
 If this is their latest  
 pretty pathetic, like  
 stupid pop-up and b  
 vaccines!



## Additional reporting

Sometimes, different news story. You can take a look sites.

## Content that you shared



SCIENCE FEEDBACK  
**False: Current measles outbreak in US do not cause**

[Learn more](#) about how fact checkers

5 h ·

Tired of fact checkers censoring posts claiming it's partly false when it's mostly true?! Follow the below mentioned cheat code. Go to settings>privacy>blocking>search for fact check/fact checkers and block them all. Since I have done this none of my stuff have been censored... I keep checking every once in a while since more popped up I have blocked like 50 so far. A way around censorship. Spread the knowledge.

Like

Share



Science Feedback

**False: US measles outbreak due to unvaccinated Americans who**

social networks, IT companies,  
**governments**, civil society (NGO,  
journalists, international agencies),  
**researchers** (academia/industry)

désinfox  
**CORONAVIRUS**

info coronavirus

Des

National Union of Journalists (SNJ):  
“infringement by the government on the  
freedom of the press”

Le vrai du faux du COVID-19. Cet  
émissaire aux articles de médias français luttant,

12 PM - Apr 30, 2020

(i)

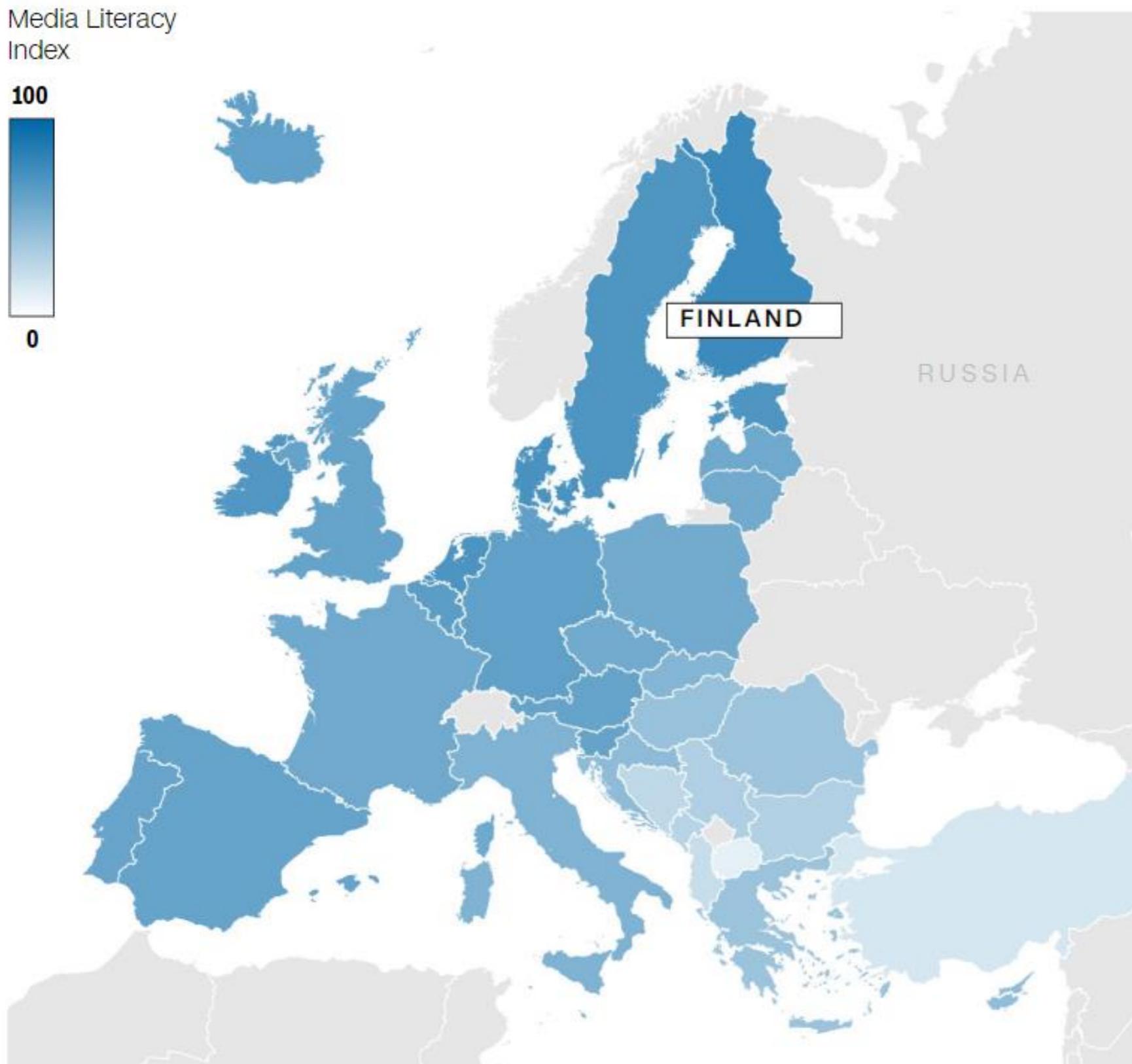
# Media literacy across Europe

Finland ranked first out of 35 countries in a study measuring resilience to the post-truth phenomenon

Media Literacy Index

100

0



Finland

Denmark

Netherlands

Sweden

Estonia

Ireland

Belgium

Germany

Iceland

UK

Slovenia

Austria

Spain

Luxembourg

Portugal

France

Latvia

Polan

Czech Rep.

Lithuania

Italy

Slovakia

Malta

Croatia

Cyprus

Hungary

Greece

Romania

Serbia

Bulgaria

Montenegro

Bosnia

Albania

76

71

70

69

69

68

64

62

62

60

60

60

59

59

56

56

55

55

55

50

48

47

44

43

40

39

38

31

30

28

25

22

- Data driven content checking is happening

- More tech
  - Ambiguity
  - Better explanations

[www.chistera.eu/projects/cimple](http://www.chistera.eu/projects/cimple)

**Interpretation 1**

Time : March 2021, April 2021  
 Country : World  
 Table : Total confirmed  
 Formula :  $(b-a)/a = 0$

According to this interpretation, the given statement seems to be **true**

The total number of confirmed cases remained constant

**Interpretation 2**

Time : 4/13/21, 4/14/21  
 Country : World  
 Table : Total confirmed  
 Formula :  $(b-a)/a = 0$

According to this interpretation, the given statement seems to be **false**

The total number of confirmed cases remained constant

- But it's not only about tech
  - Regulation
  - Education

- Harm  
**- Falsity**

- Politics  
- Resilience

