

Computational fact-checking

Problems, state of the art, and perspectives

Julien Leblay ¹ Ioana Manolescu ² Xavier Tannier ³

The Web Conference - April 2018

¹Artificial Intelligence Research Center – AIST, Japan

²Inria, LIX – CNRS and École Polytechnique, France

³Sorbonne Université – Inserm – LIMICS, France

Disclaimer

The slides presented during the WWW tutorial contained copyrighted materials.

In the present version, copyrighted materials have been replaced by blank frames.

Who are we?

Julien Leblay

Research Scientist at the
AIST's AI Research
Center in Tokyo, Japan.



Ioana Manolescu

Senior researcher at Inria
Saclay and Ecole
Polytechnique, France.



Xavier Tannier

Professor at
Sorbonne Université and
researcher at LIMICS,
France.



What is this tutorial about?

- ▶ **About** how computer science can help *a posteriori* fact checking of claims:
 - ▶ **Extracting claims** from some discourse,
 - ▶ **Searching for the facts** the claims are based on,
 - ▶ **Assessing the accuracy** of the claim,
 - ▶ **Providing perspective** to claims
- ▶ **Not only about** fake news detection!
- ▶ **Not about** image and video fact checking

Companion paper in “Journalism, Misinformation and Fact-Checking” track (**Wed, Apr 25**) and at <https://hal.archives-ouvertes.fr/hal-01722666>

Definition

news¹ *noun*

newly received or noteworthy information, especially about recent events

journalism *noun*

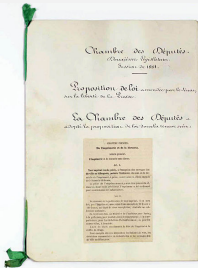
the activity of writing for newspapers, magazines, or websites or preparing news to be broadcast

- ▶ Journalists **investigate**, **check the facts**, **explain**, abiding by ethical principles including **accuracy**, **objectivity**, **impartiality**, and **accountability**
- ▶ Many countries have laws protecting freedom of the press, which also define the **rights** and **responsibilities** of news organizations.

¹ All definitions according to Oxford dictionary

Freedom of the press

In France



- ▶ The law dates back from 1881
- ▶ Born on the aftermath of insurrections in Paris ("Commune"), where **defamation** was widespread
- ▶ Already forbids publishing **fake news** causing "disturbance of the public sphere"

- ▶ France is considering stricter regulations on news during elections

Regulating the news is one thing, but where to draw the line is another.



ASIA PACIFIC

India's 'Fake News' Crackdown Crumbles Over Journalists' Outrage

Free press is an essential ingredient of a democracy

To **debate** and **express dissent**

Romania, circa 1989:



Banner reads: "Ceaușescu re-elected at the 14th congress!"
He was in power since 1965.
Massive protests lead to approx 1000 dead.
No one convicted

To **expose** and **explain** how a society functions

Panama Papers, 2016:

<https://www.icij.org/investigations/panama-papers/the-power-players/>

ICIJ INVESTIGATIONS > THE PANAMA PAPERS JOURNALISTS BLOG

Panama Papers The Power Players

Konrad Mizzi
Maltese Minister for Energy and Health (2014-present); Deputy Leader for Party Affairs, Labour Party of Malta (2016-present)

Related countries
Malta

Konrad Mizzi is the minister of energy and health of the Republic of Malta. He was elected to Parliament in 2013 and became the deputy leader of Malta's ruling Labour Party in February 2016. Given Mizzi's quick rise, one pundit called

f t

Embed

Massive tax evasion offshore.
Known thanks to work by the International Consortium of Investigative Journalism (ICIJ).

Free press is an essential ingredient of a democracy

- ▶ To **debate** and **express dissent**.
- ▶ To **confirm** or **refute** public statements.
- ▶ To **expose** and **explain** how society functions.
- ▶ To keep the authorities **accountable**.

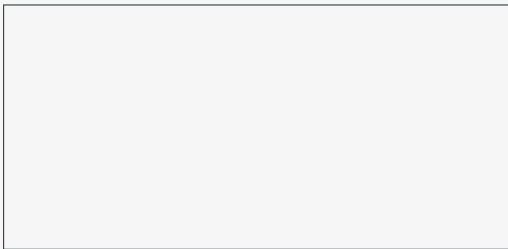
Daphne Caruana Galizia
(1967-2017)



Long standing issues

Honest mistakes

- ▶ from **incomplete** and **inaccurate** sources
- ▶ from **ambiguities** of languages



Source: <https://www.nytimes.com/2016/09/13/well/eat/how-the-sugar-industry-shifted-blame-to-fat.html>

Long standing issues (continued)

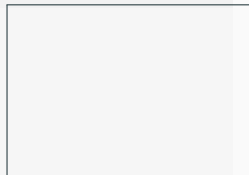
Bias

On the part of journalists and reader

- ▶ from cultural, financial or political **pressure**
- ▶ as well as other **social** or **psychological** factors

Deception (including fake news)

- ▶ as old as journalism
- ▶ can take many forms
(rumors, hoax, propaganda, satire, etc.)



Fact checking

Definition

fact-check *verb* [with object]

investigate (an issue) in order to verify the facts

Term in use since 1930 approx.



Source: Google N-gram viewer

Fact checking (Ye Good Ol' Days)

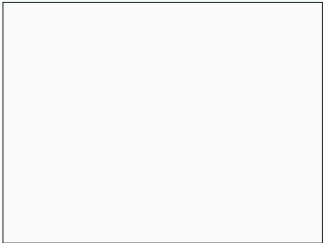
"The day I became a fact-checker at The New Yorker, I received one set of red pencils and one set of No. 2 pencils. [...] The red pencils were for underlining passages on page proofs of articles that might contain checkable facts [...] confirmed with the help of reference books from the magazine's library, including Merriam-Webster's Geographical Dictionary, the New Grove Dictionary of Music and Musicians and Burke's Peerage and Gentry."

Source:

[nytimes.com/2010/08/22/magazine/22FOB-medium-t.html](https://www.nytimes.com/2010/08/22/magazine/22FOB-medium-t.html)



Fact checking in the Internet era

- 
- ▶ As the **Internet** took off, in the mid-90s, it gradually **incorporated** all other forms of media...
 - ▶ ... allowing **anyone** publishing **anything**, while reaching a **global** audience.
 - ▶ Gradually, journalists had to become more **tech-savvy**.

Fact checking has moved from **before** to **after** publication!

- ▶ A seminal article by [?] gave birth to **computational journalism** as a discipline
- ▶ Since then, DB, IE, NLP, ML, KR communities have started work in the area

Context and problems

Q: "Is it true that in Moscow, Mercedes cars are being given to citizens?"

A: "Yes, but it is not Moscow but Leningrad, not Mercedes but Ladas, and not given to but stolen from."

Yeravan jokes, famous in the Eastern block during communism.

Fact-checking in the Internet era: what's new ?

The Web as the primary media

- ▶ Traditional news sources increasingly disseminate through the Web
- ▶ New outlets, e.g. so-called pure players, runs 100% of their operations on the Web
- ▶ Social networks became major media outlets and conduit

Fact-checking in the Internet era: what's new ?

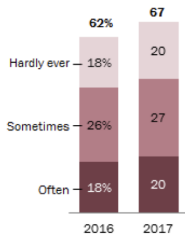
“Democratization” of authorship

- ▶ **Non-media** organizations (companies, government Web sites) and individuals gained access to large scale **publishing** facility
- ▶ **No** editorial process or **ethics** required
- ▶ Line **blurred** between news **producers** and **consumers**.
- ▶ Sudden **abundance** of data (with varying quality/credibility)

Social networks have become a primary source for news

In 2017, two-thirds of U.S. adults get news from social media

% of U.S. adults who get news from social media sites ...

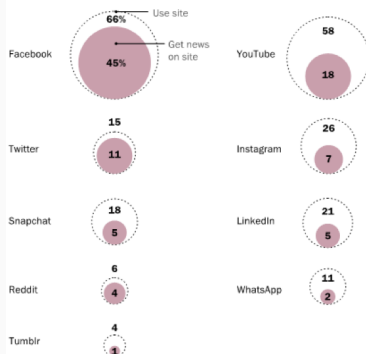


Source: Survey conducted Aug. 8-21, 2017.
"News Use Across Social Media Platforms 2017"

PEW RESEARCH CENTER

Social media sites as pathways to news

% of U.S. adults who use each social media site and % of U.S. adults who get news from each site

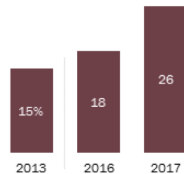


Source: Survey conducted Aug. 8-21, 2017.
"News Use Across Social Media Platforms 2017"

PEW RESEARCH CENTER

About one-in-four now get news from multiple social media sites

% of U.S. adults who get news on two or more different social media sites



Source: Survey conducted Aug. 8-21, 2017.
"News Use Across Social Media Platforms 2017"

PEW RESEARCH CENTER

In some emerging countries, Internet "is" Facebook ²

²qz.com/243086/facebook-is-creating-a-parallel-internet-in-emerging-markets/

Social media mishaps

Social networks are increasingly weaponized to spread dubious information

Example (Recent events)

- ▶ 2-3% of Facebook accounts are fake, 5% on Twitter ^a
- ▶ Twitter conducted a sweeping “bot purge” in February 2018 ^b
- ▶ Russian meddling in the US presidential election
- ▶ Cambridge Analytica scandal vs. US elections and Brexit
- ▶ Insane man’s terror act in Germany wrongly connected to immigrants^c

^a[nytimes.com/2017/11/03/technology/facebook-fake-accounts.html](https://www.nytimes.com/2017/11/03/technology/facebook-fake-accounts.html)

^b[thedailybeast.com/inside-twitters-bot-purge](https://www.thedailybeast.com/inside-twitters-bot-purge)

^c[lemonde.fr/international/article/2018/04/09/l-allemande-sous-le-choc-apres-l-attaque-de-munster-et-l-attentat-dejoue-a-berlin_5282856_3210.html](https://www.lemonde.fr/international/article/2018/04/09/l-allemande-sous-le-choc-apres-l-attaque-de-munster-et-l-attentat-dejoue-a-berlin_5282856_3210.html)

A brief history of fact checking initiatives



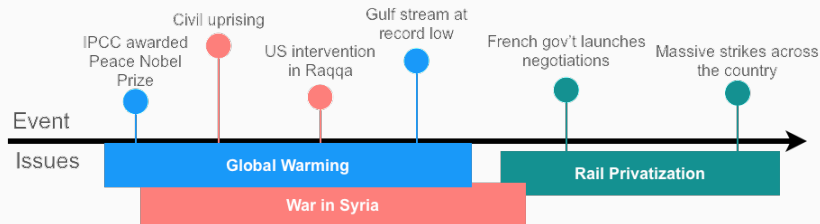
Fact checking sites today



Source: reporterslab.org/fact-checking (Duke Reporters Lab)

Types of information

World **events** are intertwined with **longer-term** social **issues**.



- ▶ Information **is altered** as it propagates across a social network (through bias, accumulated errors, and outright lies)
- ▶ Journalists must provide a **short** and **high quality channel** between the events on the public
- ▶ Stakeholders can be **motivated**, rely on **rhetoric**, **persuasion**
- ▶ Journalists must **balance emotions**❤️, **credibility**👤 and **reasoning**🧩.

Computer science and journalism: how can we help?

1. **Data journalism**: journalistic work significantly or mainly based on (digital) data
2. **(Semi)-automated fact checking**: fact checking work where some tasks are delegated to software
 - ▶ Our focus today
 - ▶ Fact checking tasks will be detailed shortly
3. **Fake news detection**: software which estimates the level of falsehood of a piece of news
 - ▶ True, false, in-between...
 - ▶ May not use reference sources.

Fact checking ingredients

To successfully check a claim, one needs to:

1. Lift the **ambiguity**

✖ Vague statements lead to too many distinct interpretations, which one to check?

✖ Clarify the context in which the claim is analysed (space, time...).

2. Ensure it is **backed** by **sufficient** references to **sources**.

📌 **Reliable reference sources** give the background against which to check.

3. Validate the claim as **consistent** with the sources.

✖ Some claims are **crafted** to mislead, i.e., look valid wrt a context or source that is irrelevant or flawed.

The need for transparency

The International Fact-Checking Network (IFCN) is sponsored by the Poynter Institute to “promote excellence in fact checking”.

Members commit to:



1. Non-partisanship and fairness.
2. Transparency of sources.
3. Transparency of function and organization.
4. Transparency of methodology.
5. Open and honest corrections.

Source: poynter.org/international-fact-checking-network-fact-checkers-code-principles

The limits of fact checking

- ▶ **Confirmation bias:** people are more likely to believe what fits their prior views.
 - ▶ Man-made part of the **echo chamber**.
 - ▶ Automated recommendation systems trap users in **filter bubbles**.

Yet:

- ▶ Filter bubbles and echo chambers are still being studied [?, ?].
- ▶ Showing readers links to “**related stories**” **reduces misperceptions** more effectively [?].

The limits of fact checking: Timing matters!

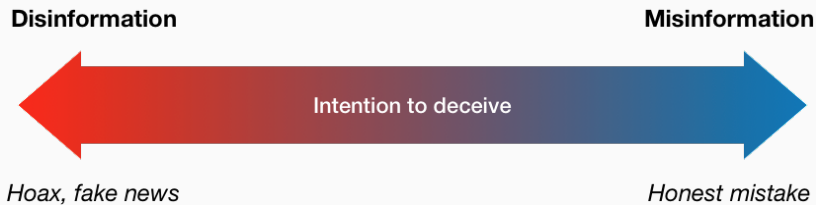
Emotionally engaging information, such as **rumors** and **propaganda**, spread **faster** than **corrections** on social networks [?].

- ▶ False news spread faster than true ones; most of the audience is reached in the first 24 hours [?].
- ▶ If **verification** comes **too late**, false information has time to “**stick**” with audience³.
- ▶ **Backfire effect**: defiance towards fact checkers may reinforce reader’s perception if confronted directly [?], **near instant**-correction making things worse [?].



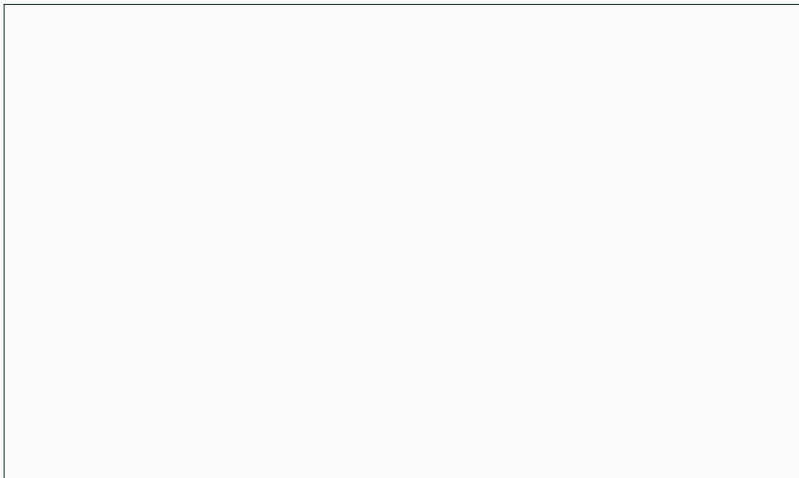
³jonathanstray.com/networked-propaganda-and-counter-propaganda

Misinformation and disinformation



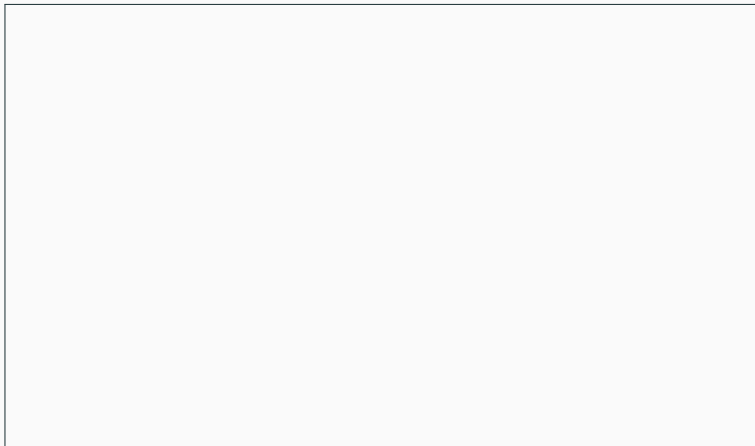
Rumors, myths, conspiracy theories

Widespread, often long-standing, misconceptions, on which one might base their judgment.



Hoax

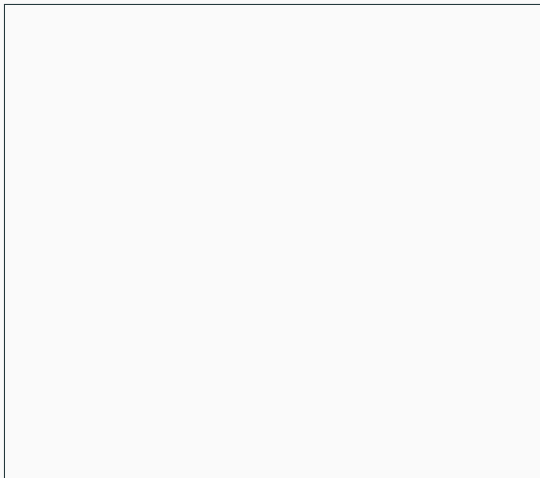
Non-elaborate, unsubstantiated claim. Aimed at spreading virally.



Source: snopes.com/fact-check/david-hogg-on-campus-rumor-hoax/

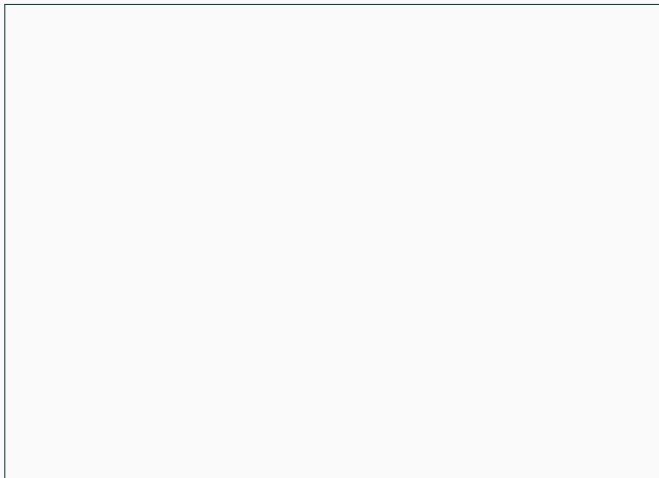
Clickbait

Catchy title, poor content. Aimed mainly at attracting audience for financial gains.



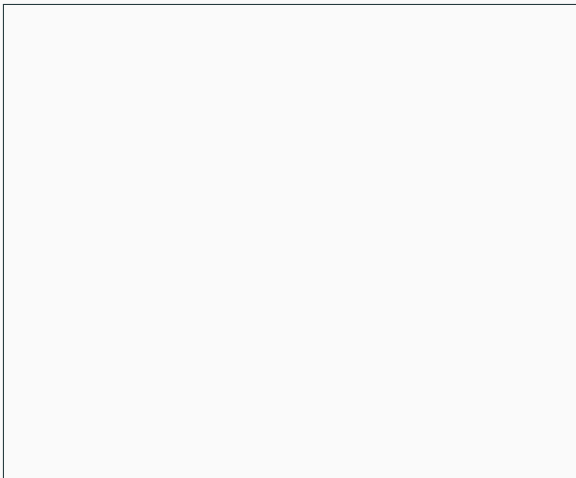
Media hype

Catchy story, with some core element of truth, but vastly exaggerated.
Frequent in health, technology and science news.



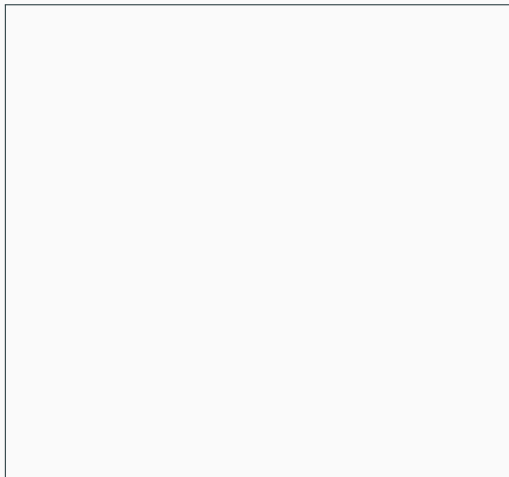
Wrong context assignment

Genuine content (image, video, audio, or quote) planted in unrelated context to steer opinion.



Content doctoring

Realistic depiction of events that did not occur.



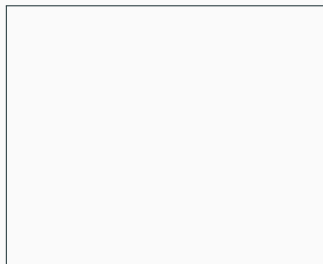
Incorrect factual claims

Claims with **obvious interpretation** and for which **there exists** reasonably **relevant** and **accurate data**.

“Our prisons are filled-up with foreigners.”

BBC Question Time audience member, Oct. 20, 2016

Foreign citizens make up 9% of the general population and 12% of the prison population in England and Wales. [...] The number and proportion of foreign prisoners is falling: there were over 11,000 foreign prisoners in 2010.

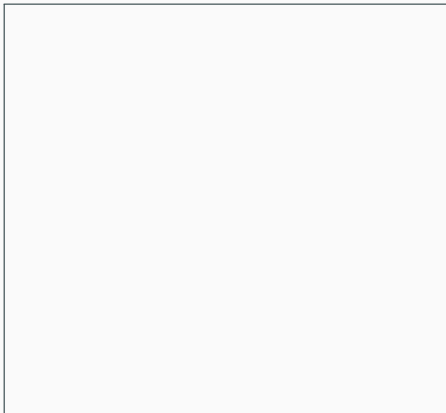


Source: fullfact.org/immigration/foreigners-prison

Ambiguous or oversimplifying claims

The claim is open to **multiple interpretations**, some of which may be true, but not necessarily the most relevant one.

Typically **requires** in-depth **analysis**.



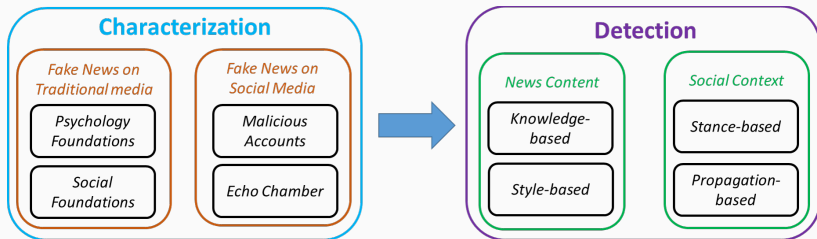
Flavours of fake news [?]

Even when there is intention to deceive, the **purpose** of the deception may vary a lot:

- ▶ **coordinated** and **well-targeted** information forgery
- ▶ **simple lies** that catch on to a large audience
- ▶ **humor**, satire, sarcasm

From characterization to detection [?]

Fake news is a news article that is intentionally and verifiably false.

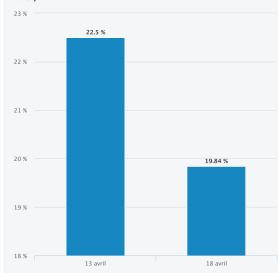


Correct but imprecise

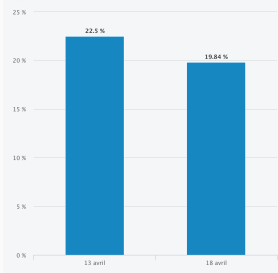
By Les décodeurs of LeMonde.fr news papers.

- ▶ The French Railway Company (SNCF) is undergoing strike following reforms announcement.
- ▶ Unions, the company, the government and other interest group are in show of force.
- ▶ The SNCF recently published a press release aluding the mobilization was falling.

Le taux de participation à la grève du 18 avril (avec une valeur minimum à 18 %)



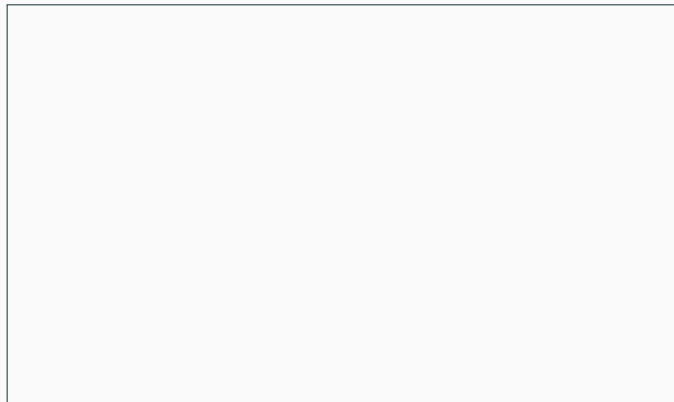
Le taux de participation à la grève du 18 avril (avec une valeur minimum à 0)



Source: lemonde.fr/les-decodeurs/article/2018/04/18/le-graphique-trompeur-de-la-

In-depth analysis

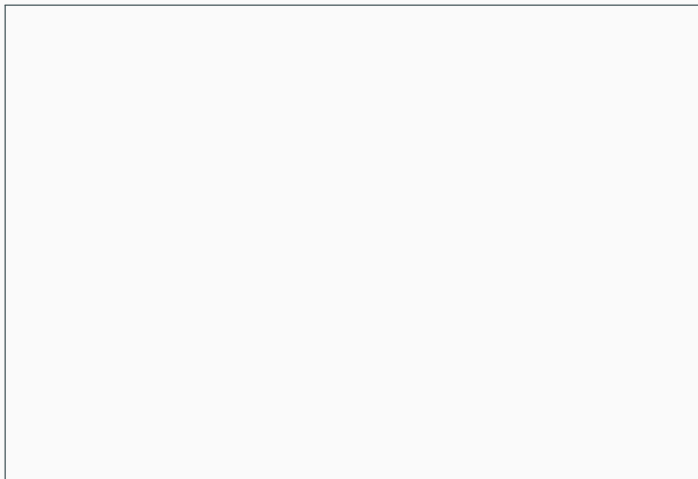
Detailed “forensic” study of **past events** for which reference sources were hard to obtain, witnesses speak late etc.



Source: <https://www.mediapart.fr/journal/france/040418/nicolas-sarkozy-bien-servi-les-interets-de-kadhafi-voici-les-preuves>

Promise verification

Validating **past** claims made about **the future**

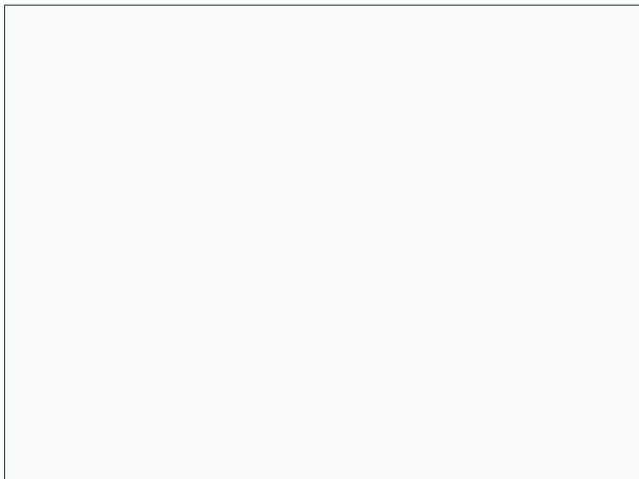


Source:

politifact.com/truth-o-meter/promises/obameter/subjects/politifacts-top-promises/

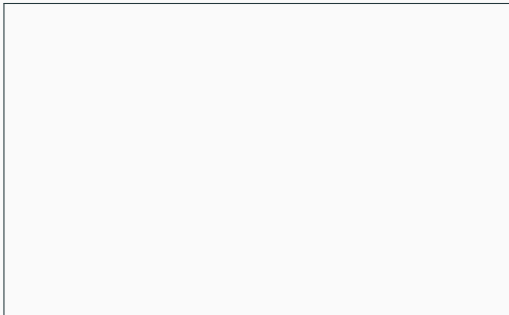
Reversal tracking

Checking a personality's **position** or **stance** on a specific issue **over time**.

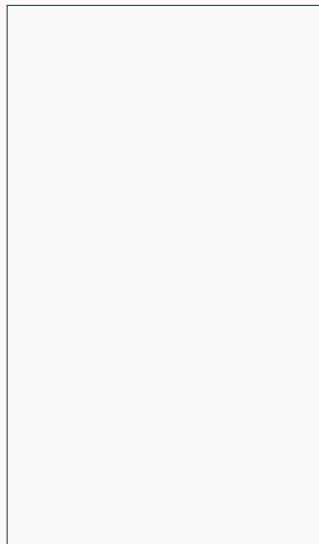


Source: politifact.com/texas/statements/2018/jan/10/beto-orourke/beto-orourke-flip-flops-requiring-public-service-y/

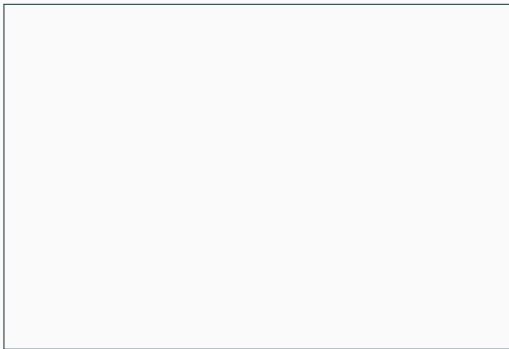
State of the art



- ▶ In-depth analysis of a claim
- ▶ First politics, now science, health etc.
- ▶ Particularly active during US elections

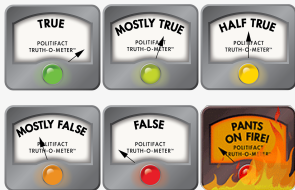


“We are not going to let our campaign be dictated by fact-checkers.”



- ▶ In-depth political claim analysis
- ▶ Simple classification for checked claims
- ▶ Position reversals
- ▶ DB published through an API

Truth-o-meter



Flip-o-meter



Désintox blog from “Libération”

Crosscheck from First Draft News

- ▶ Supported by Google News Initiative
- ▶ Relies on volunteers
- ▶ Trains the public to critical thinking and news analysis

Fact checking pipeline

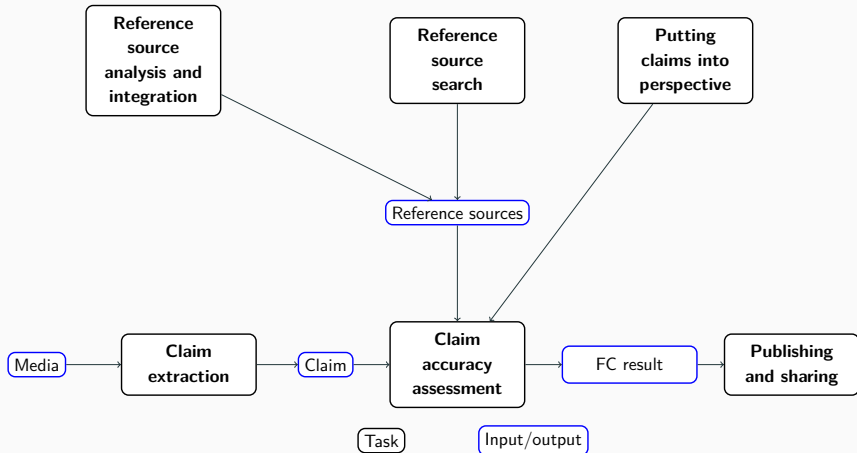
Definition (Fact checking [?])

Defined as a four-stage process where

- (i) media sources are **monitored**,
- (ii) claims are **spotted**,
- (iii) claims are **checked**,
- (iv) fact checking analysis **results** are created and **published**.

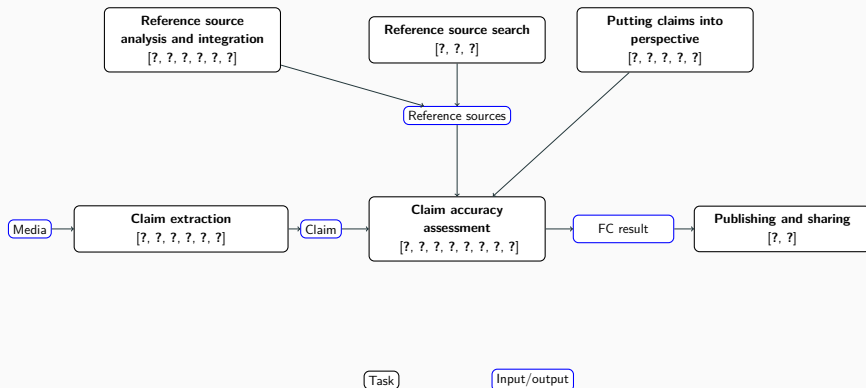
Fact checking from a content management perspective

Overview of **tasks**, **inputs** and **outputs**



Fact checking from a content management perspective: partial reference list

Full list of references available in companion paper [?] (presented **Wednesday!**)



Which reference data source to use?

► Fixed (known in advance):

- **ClaimBuster** [?]
- **DisputeFinder** (PolitiFact API) [?]
- **FullFact** (internal DB of manually checked claims) [?]
- **TruthTeller** (claims manually checked by Factcheck.org);
- The **Décodex** plug-in developed by Le Monde also leverages their past fact checking analyses

► Web search:

- **DeFacto** [?, ?]
- **ClaimBuster** [?]

Professional journalism is **very** picky on source quality.



Which reference data source to use? (continued)

- ▶ **General Knowledge Bases** such as **Wikipedia** [?]
- ▶ **Heterogeneous Open Data** **FactMinder** manual analysis of online articles against any open data sources [?].



the star

COMMENTARY

 HOME

 YOUR TORONTO

 OPINION

 Commentary

 Letters

 Public Editor

 Editorial Board

 Contact Us

 SPORTS

 BUSINESS

 ARTS & LIFE

 OPINION

 CLASSIFIEDS

 OBITUARIES

 HOME

 YOUR TORONTO

 OPINION

 Commentary

 Letters

 Public Editor

 Editorial Board

 Contact Us

 SPORTS

 BUSINESS

 ARTS & LIFE

 OPINION

 CLASSIFIEDS

 OBITUARIES

 HOME

 YOUR TORONTO

 OPINION

 Commentary

 Letters

 Public Editor

 Editorial Board

 Contact Us

 SPORTS

 BUSINESS

 ARTS & LIFE

 OPINION

 CLASSIFIEDS

 OBITUARIES

 HOME

 YOUR TORONTO

 OPINION

 Commentary

 Letters

 Public Editor

 Editorial Board

 Contact Us

 SPORTS

 BUSINESS

 ARTS & LIFE

 OPINION

 CLASSIFIEDS

 OBITUARIES

 HOME

 YOUR TORONTO

 OPINION

 Commentary

 Letters

 Public Editor

 Editorial Board

 Contact Us

 SPORTS

 BUSINESS

 ARTS & LIFE

 OPINION

 CLASSIFIEDS

 OBITUARIES

 HOME

 YOUR TORONTO

 OPINION

 Commentary

 Letters

 Public Editor

 Editorial Board

 Contact Us

 SPORTS

 BUSINESS

 ARTS & LIFE

 OPINION

 CLASSIFIEDS

 OBITUARIES

 HOME

 YOUR TORONTO

 OPINION

 Commentary

 Letters

 Public Editor

 Editorial Board

 Contact Us

 SPORTS

 BUSINESS

 ARTS & LIFE

 OPINION

 CLASSIFIEDS

 OBITUARIES

Dashboard		Annotator			
People <ul style="list-style-type: none"> Justin Trudeau Jason Kenney Barack Obama 	Places <ul style="list-style-type: none"> Washington London Toronto Ottawa United States Canada 	Countries <ul style="list-style-type: none"> United States Canada 			
Organizations <ul style="list-style-type: none"> Hisper government 		Cities <ul style="list-style-type: none"> Washington London Toronto Ottawa 			
Who said what ?		Bio			
Barack Obama <p> Its fight against terrorism has succeeded in averting mass terror attacks on U.S. soil – but at a very steep cost </p> <p> after 11 years of wartime footing, a new approach is required. </p> <p> for the U.S. to define [a] effort not as a boundless "global war on terror," but rather as a series of persistent, targeted efforts to dismantle specific networks of violent extremists that threaten America. </p>		<p> Barack Hussein Obama II is the 44th and current President of the United States. He is the first African American to hold the office. In January 2009, Obama was sworn in as a U.S. Senator in the state of Illinois. He would hold this office until November 2008, when he resigned following his victory in the 2008 presidential election. Born in Honolulu, Hawaii, Obama is a graduate of Columbia University and Harvard Law School, where he was the president of the Harvard Law Review. </p>			

Building reference data sources: truth discovery

Partially overlapping Web sources require **arbitrating** between their information.

Example

NY restaurant information [?]

Source	Coverage	Exactness	Freshness	#Closed-rest
MenuPages	.66	.98	.86	29
TasteSpace	.44	.97	.3	106
NYMagazine	.43	.98	.54	59
NYTimes	.43	.98	.38	72
ActiveDiner	.41	.95	.86	70
TimeOut	.38	.99	.68	33
SavoryCities	.27	.99	.41	33
VillageVoice	.22	.94	.4	37
FoodBuzz	.18	.92	.3	59
NewYork	.13	.92	.45	28
OpenTable	.12	.92	.45	9
DiningGuide	.1	.9	.09	48
GoogleMaps	-	-	-	212

Extracted from 12 sources + manually checked

Probabilistic approach for determining the true value, based on coverage, exactness and freshness, and on who copied whom.

Building reference data sources: truth discovery

Truth discovery survey [?]:

Input: a set of **values** for an **object**, each from a different **source**

Output: **most likely value** and **trustworthiness** of each source

Principle:

- ▶ A source whose value for an object was deemed correct, will be considered more trustworthy
- ▶ ... and values coming from a trustworthy source will be considered more likely to be correct

Methods: iterative; optimization-based (error minimization);
probabilistic graphical models

Constructing reference data sources: data integration

Valuable information is sometimes found **across** several data sources

Data integration approaches:

- ▶ **Warehouse**: extract and consolidate all data sources into one



Text (contracts) and relational (screen company coordinates) data sources fused into one (Neo4J) graph database

Easy to use; needs to be redone for every new dataset

- ▶ **Mediator**: structured data sources remain unchanged and are queried together under a **unified schema**

Constructing reference data sources: data integration (cont'd)

Valuable information is sometimes found **across** several data sources

Data integration approaches (cont'd):

- ▶ **Data space**: structured and unstructured data sources queried through keywords [?]
- ▶ **Data lake**: large number of structured and unstructured data sources w/o unified schema; subsets of these are exploited together in mediator style, e.g. [?]⁴
- ▶ **Dataflow**: data journalism analytical pipelines⁵

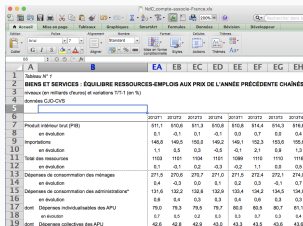
⁴<https://www.ibm.com/analytics/data-management/data-lake>,
<https://blogs.oracle.com/bigdata/the-new-data-lake-you-need-more-than-hdfs>

⁵<http://jonathanstray.com/introducing-the-cj-workbench>

Improving usability of reference data sources

High-quality reference data, e.g., published by statistic institutes, may be hard to query

1. **Extract** data into RDF Linked Open Data (preserving table and header structure) [?]
2. **Search** for exact or closest approximate answer to a keyword query [?]



	2010	2011	2012	2013	2014	2015	2016	2017	2018
Produit intérieur brut (PIB)	511,3	510,8	511,3	510,8	510,8	514,8	515,3	516,8	517,3
en évolution	0,1	-0,1	0,1	-0,1	0,0	0,7	0,9	0,4	0,4
Importations	140,0	140,5	140,8	140,2	140,1	142,3	143,6	145,6	146,6
en évolution	1,1	0,3	0,3	-0,5	-0,1	2,1	0,9	1,3	0,7
Total des ressources	1103	1101	1104	1101	1099	1116	1110	1116	1118
en évolution	0,1	-0,1	0,2	-0,3	-0,2	1,1	0,9	0,5	0,5
Dépenses de consommation des ménages	271,8	270,6	270,7	271,0	271,5	272,4	272,1	274,0	274,0
en évolution	0,4	-0,3	0,0	0,1	0,2	0,3	-0,1	0,7	0,0
Dépenses de consommation des administrations*	131,0	132,2	132,8	132,9	133,4	134,3	134,5	134,8	134,8
en évolution	0,8	0,4	0,5	0,3	0,4	0,6	0,3	0,3	0,0
Total Dépenses individuelles des APU	79,0	79,3	79,5	79,7	80,2	80,5	80,7	81,1	81,1
en évolution	0,0	0,2	0,2	0,2	0,3	0,3	0,2	0,4	0,0
Total Dépenses collectives des APU	42,6	42,8	42,9	43,0	43,3	43,5	43,6	43,6	43,6

Source: insee.fr/fr/statistiques/3292347?sommaire=3292415

Searching for truth in statistic tables

Query: “youth unemployment France August 2017”

	Seasonally adjusted youth (under 25s) unemployment				
	Number of persons (in thousands)				
	Oct-2016	Jul-2017	Aug-2017	Sep-2017	Oct-2017
Belgium	:	77	77	77	:
Bulgaria	27	22	21	19	19
Czech Republic	34	27	25	23	23
Denmark	62	54	53	49	47
Germany	293	283	283	283	283
France	663	629	625	623	625

Answer: 625 + link to the spreadsheet as result proof (provenance, justification)

- ▶ **Extraction** needs to cope with nested headers
- ▶ Off-line source **indexing**
- ▶ Search for (i) relevant **datasets** and (ii) most relevant **cells** in each dataset

Claim recognition: claim extraction

- ▶ Topic-driven extraction from media articles [?].

- ▶ Task: Given a topic (*context*), find related claims, e.g.:

Topic

Selling violent video games to minors should be banned

Related claim

Violent video games can increase children's aggression

- ▶ Approach: fully supervised learning.

- ▶ Disputed claims covered by the reference database:


DisputeFinder [?].

- ▶ Task: Given a text, extract claims disputed by a trusted source, e.g.:

Many **vaccines contain mercury, aluminium and other toxins** that should have parents asking questions before immunizing their children.

- ▶ Approach: keyword retrieval against a claim database.

Claim recognition: claim extraction (continued)

- ▶ Entity disambiguation applied on claims, using a reference knowledge base: **DeFacto** [?].
 - ▶ Task: Given a text, extract 10 types of predefined relations between named entities.⁶
 - ▶ Example:
Input: Albert Einstein was awarded the Nobel Prize in Physics.
Output:
dbr:Albert_Einstein dbo:award dbr:Nobel_Prize 
 - ▶ Approach: rule-based
- ▶ Research into extraction from text feeds, audio, video: **FullFact** [?]
(technical details not available at this time).

⁶Relations are: award, birth, death, foundationPlace, leader, NBAteam, publicationDate, spouse, starringActor, subsidiary

Claim recognition: classifying check-worthiness

1. Verifiability: **Verifiable** vs. **Unverifiable** [?, ?, ?]
2. Factuality and worthiness:
Non-factual (e.g., opinions or subjective content) vs.
Factual but **not interesting** (consensual, general) vs.
Factual and **interesting** (that is, check-worthy). [?, ?]
3. Opinion: **Facts** vs. **opinions** [?]
4. Dialogic and argumentative markers:
 - ▶ Degrees of **agreement** with a previous post
 - ▶ Cordiality, audience-direction, combativeness, assertiveness, emotionality of argumentation, sarcasm[?]

All these approaches based on fully-supervised systems with expert- or crowd-sourced data.

Stance detection

Is a text **in favor** of a given target, **against** it, **neutral** or **unrelated**?

- ▶ Target: legalization of abortion
- ▶ Negative stance: *"A foetus has rights too! Make your voice heard"*.
- ▶ Target: Donald Trump
- ▶ Positive stance: *"@realDonaldTrump is the only honest voice of the @GOP"*.

Sources can be **general claims**, **debates** in online forums, student **essays**, but mostly news or political speeches, debates, tweets.

Approaches are all based on supervised learning.

[?, ?, ?, ?, ?, ?, ?, ?, ?, ?], FakeNewsChallenge (2017).

Claim accuracy assessment

- ▶ **DeFacto** finds evidence potentially proving the claim as Web page text snippets, sufficiently close to the claim [?, ?].
 - ▶ **ClaimBuster**
 - 1) Matches claims against previously checked claims from trusted repositories (PolitiFact, etc.)
 - 2) If no similar claim is found, reverts to **Web search** engines and **question answering** systems such as **Wolfram Alpha**.
- [?, ?]

In [?], checking an RDF triple claim is finding a **short path** from the subject to the object, in the transitive closure of a knowledge graph. The **degree** of the node influences the **truth value** of a candidate supporting path.



Claim accuracy assessment (continued)

- ▶ The **Fast and Furious FactCheck Challenge**⁷ proposed to classify news articles (not claims) among: TRUE, FALSE, SOMEWHAT TRUE and SOMEWHAT FALSE w/ human and/or automated tools;
- ▶ Les **Décodeurs**⁸ (Le Monde) developed:
 - ▶ A database of manually checked claims w/ analysis and rumor propagators.
 - ▶ A web navigator plugin w/ a trust score from the aggregated outputs of previous fact checks, where available.

⁷<https://herox.com/factcheck/>

⁸<http://www.lemonde.fr/les-decodeurs/>

Claim accuracy assessment: related tasks

These well-known NLP tasks have never really been applied to fact-checking problems as such:

- ▶ **Textual entailment** compares two texts and decides whether one implies the other [?].
- ▶ The **SemEval**'s Semantic Textual Similarity task offers a graded and typed definition of semantic similarity [?].
- ▶ **Rumor detection** classifies a set of posts/tweets as rumor or not rumor, or studies the birth and propagation of rumors [?, ?].

Fake news detection

- ▶ Flourishing field
- ▶ Growing number of challenges, hackathons and data sets available
 - ▶ Around 160 news-related datasets and 70 public kernels on **Kaggle**
 - ▶ **BuzzFeedNews**⁹: Sample of news published on Facebook prior to the 2016 U.S. elections
 - ▶ **LIAR**¹⁰: A Politifact archive
 - ▶ **BS Detector**¹¹: data collected through the BS detector browser extension.
 - ▶ **CREDBANK**¹²: A Large-scale Social Media Corpus With Associated Credibility Annotations

⁹<https://github.com/BuzzFeedNews/2016-10-facebook-fact-check/tree/master/data>

¹⁰https://www.cs.ucsb.edu/~william/data/liar_dataset.zip

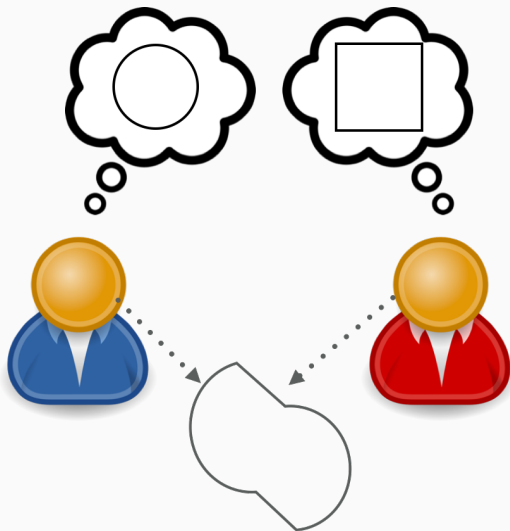
¹¹<https://github.com/bs-detector/bs-detector>

¹²<https://github.com/compsocial/CREDBANK-data>



- ▶ “Automatic Identification and Verification of Claims in Political Debates”
- ▶ **Task 1: Check-worthiness.** Predict which claim in a political debate should be prioritized for fact-checking.
- ▶ **Task 2: Factuality.** Checking the factuality of the identified worth-checking claims.
- ▶ <http://alt.qcri.org/clef2018-factcheck/index.php?id=overview>

Putting claims into perspective

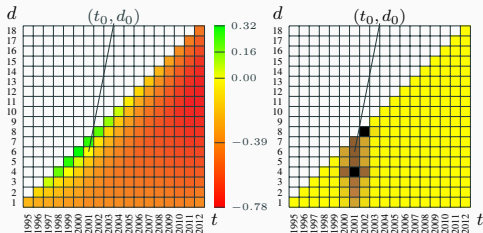


The query perturbation approach [?, ?]

Giuliani's claim: "Adoptions went up 65 to 70 percent when [he] was mayor [of New York City]."

```
SELECT after.total / before.total
FROM (SELECT SUM(number) AS total FROM adopt
      WHERE year BETWEEN t-w-d+1 AND t-d) AS before,
      (SELECT SUM(number) AS total FROM adopt
      WHERE year BETWEEN t-w+1 AND t) AS after;
```

Query Response Surface (QRS)



(a) Relative strength of results

(b) Sensibility of parameter settings

The query perturbation approach (continued)

Relative strength and relative sensitivity are used to

- ▶ Find counter-argument (that weakens the original claim), and reverse-engineer vague claims
- ▶ Robustness: All perturbations result in stronger or equally strong claims
- ▶ Other notions such as fairness, and uniqueness

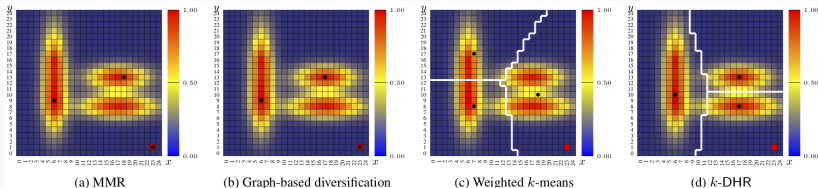
Also introduce ways to check window aggregate comparison claims, and time series similarities claims.

Diversity and representativity [?]

Follow-up work: when many counter arguments exist, select a subset maximizing utility, diversity and representativeness.

Problem: Find a Diverse Set of k High-Value Representatives from numerical data, for counter-argument generation and computational lead finding [?].

Three interesting areas (plus one noisy spike) are hidden in the data. The first three methods fail to find them and/or to ignore the spike.



Optimization method to automatically select k -DHR [?].

Putting claims into perspective (continued)

Context-dependent reasoning can be used to a **veracity** score to all possible **contexts** of a claim [?]

Example

“John Doe is a Eurosceptic.”

- ▶ Depends on what we mean by “Eurosceptic”
- ▶ Not everybody agrees!

Key idea: annotate the data and axioms with contextual details.

Axioms

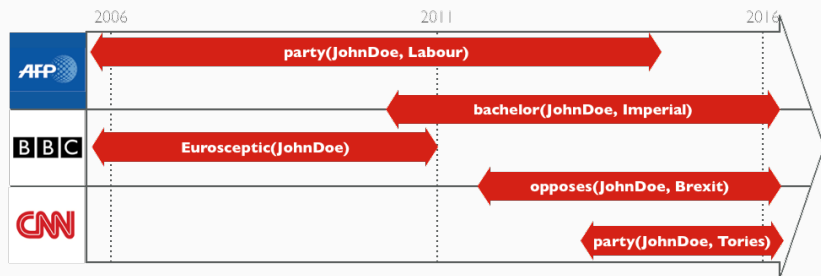
$\text{bachelorFrom}(X, Y) \rightarrow \text{CollegeGrad}(X)$	$:\{[[-\infty, +\infty], \{\text{bbc}, \text{cnn}\}]\}$
$\text{party}(X, \text{Tories}) \rightarrow \text{Eurosceptic}(X)$	$:\{[2007, 2013], \{\text{afp}, \text{bbc}\}]\}$
$\text{party}(X, Y) \wedge \text{opposes}(Y, Z) \rightarrow \text{opposes}(X, Z)$	$:\{[[-\infty, +\infty], \{\text{afp}, \text{bbc}\}]\}$
$\text{supports}(X, Y) \wedge \text{opposes}(X, Y) \rightarrow \perp$	$:\{[[-\infty, +\infty], \{\text{afp}, \text{bbc}, \text{cnn}\}]\}$

$\text{CollegeGrad}(X) \rightarrow \text{opposes}(X, \text{Brexit})$:6.0
$\text{party}(X, \text{Tories}) \rightarrow \text{support}(X, \text{Brexit})$:3.0

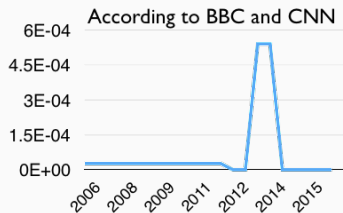
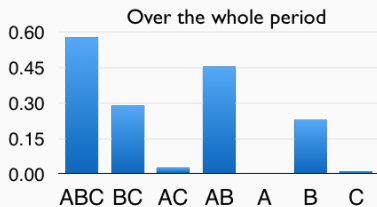
Hard

Soft

Putting claims into perspective (continued)

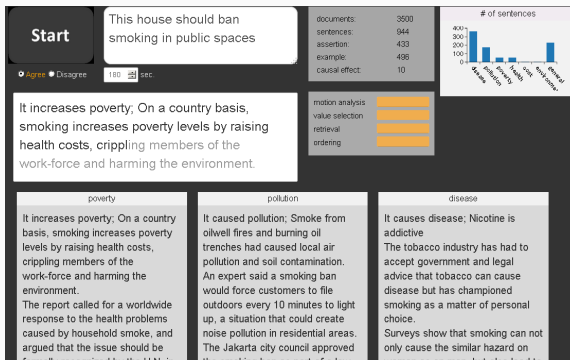


Eurosceptic(JohnDoe)?



Putting claims into perspective (continued)

- ▶ Data mining techniques used to highlight claims about behavior in voting or rating contexts:
Identify groups of individuals and situations where their agreement significantly differs from usual [?].
- ▶ Attempts to build a **general** and **balanced** picture of a complex issue [?].

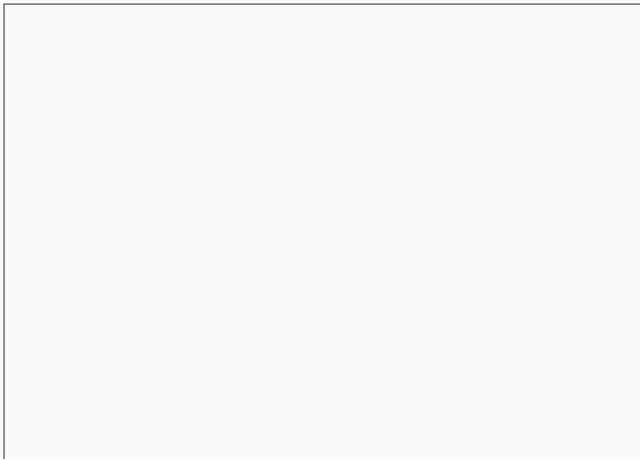


Sharing and publishing fact checking results

- ▶ **DeFacto** shares outputs as RDF graphs with provenance information [?];
- ▶ **ClaimBuster** provides access to their fact checking outputs [?, ?];
- ▶ **FactCheck.org** and **PolitFact** provide API access, and their output is already used by several other tools

Structured Journalism

- **Structured Journalism**¹³ encourages journalists to publish database items to simplify aggregating, mashing and referencing stories

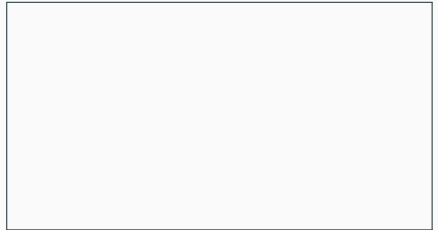
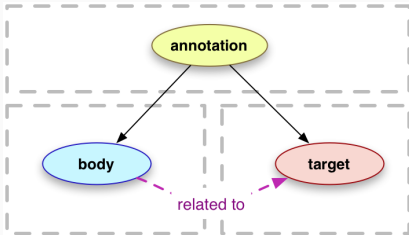


Source: <https://project.wnyc.org/traffic-deaths-2015/>

¹³<https://reporterslab.org/structured-journalism/>

Web Annotations

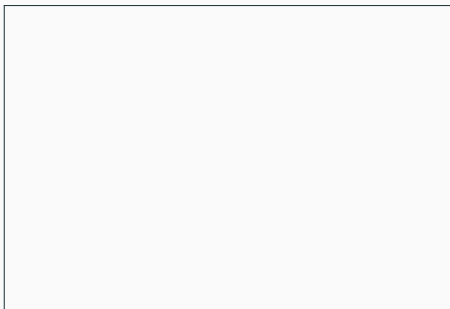
- W3C's Web Annotation Working Group published recommendation's data model, vocabulary and protocol



Source: w3.org/TR/annotation-model/

Source: web.hypothes.is/journalism/

- ▶ **ClaimReview**¹⁴ was introduced to Schema.org in 2017.
- ▶ Used by search engines to quickly find analysis on past claims.
- ▶ **Share the facts** by **Jigsaw**, an Alphabet innovation incubator, and **Duke Reporters' Lab** facilitates sharing fact checking articles.



Source: www.sharethefacts.org/

¹⁴<http://schema.org/ClaimReview>

Publishing ClaimReview using MicroFormat

```
<div itemscope="" itemtype="http://schema.org/ClaimReview">
```

```
  An example paragraph reviewing a claim expressed in another document.
```

```
  <dl>
```

```
    <dt>Date published:</dt>
```

```
    <dd itemprop="datePublished">2014-07-23</dd>
```

```
    <dt>Review url:</dt>
```

```
    <dd itemprop="url">http://www.politifact.com/texas/statements/2014/jul/23/rick-perry/
```

```
    <dt>Review by:</dt>
```

```
    <dd>
```

```
      <span itemprop="author" itemscope="" itemtype="http://schema.org/Organization">
```

```
        <span itemprop="name"><a itemprop="url" href="http://www.politifact.com/">Politi
```

```
        
```

```
      </span>
```

```
    </dd>
```

```
</dl>
```

```
<h3>Claim reviewed:</h3>
```

```
  <blockquote itemprop="claimReviewed">
```

```
    More than 3,000 homicides were committed by 'illegal aliens' over the past six years.
```

```
  </blockquote>
```

```
  <span itemprop="reviewRating" itemscope="" itemtype="http://schema.org/Rating">
```

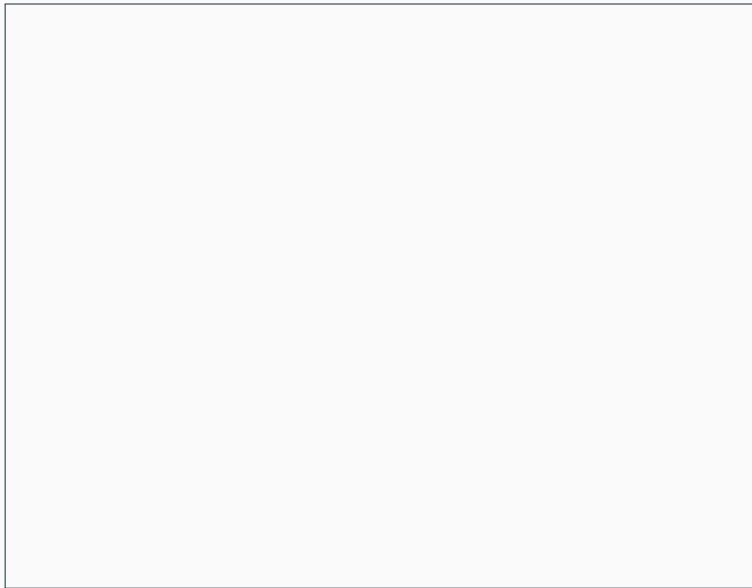
```
    Rating: <span itemprop="ratingValue">1</span>
```

```
    (best score: <span itemprop="bestRating">6</span>),
```

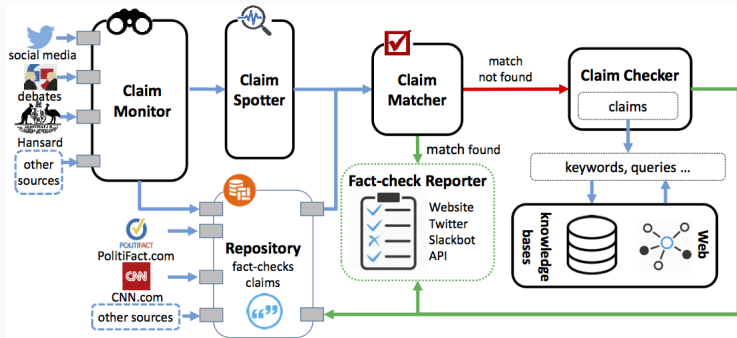
```
    "<span itemprop="alternateName">True</span>".
```

```
    <img itemprop="image" src="http://static.politifact.com.s3.amazonaws.com/rulings/to
```

ClaimReview as used in Google News



End-to-end systems: ClaimBuster [?]



- ▶ Assigns a **score** to each sentence based on how factual it is (low = subjective or opinionated phrase.)
- ▶ Above a certain **threshold**, a claim is **matched** against a set of fact-checking websites
- ▶ **Complementary evidence** collected from general knowledge bases, otherwise search engines

End-to-end system: ClaimBuster (continued)

2016 Third Presidential Debate. Oct. 19, 2016, 9 p.m. EST

Chronological Order Order by Score

Least Check-worthy >=0.1>=0.2>=0.3>=0.4>=0.5>=0.6>=0.7>=0.8>=0.9>=1.0 Most Check-worthy

is tougher. But they know what's going on. They know it better than anybody. They want strong borders. They feel we have to have strong borders. I was up in New Hampshire the other day. **The biggest complaint they have -- it's with all of the problems going on in the world, many of the problems caused by Hillary Clinton and by Barack Obama.** All of the problems -- the single biggest problem is heroin that pours across our southern border. It's just pouring and destroying their youth. It's poisoning the blood of their youth and plenty of other people. We have to **Fact-check this** ders. We have to keep the drugs out of our country. We are -- right now, we're getting the drugs, they're getting the cash. We need strong borders. We need absolute -- we cannot give amnesty. Now, I want to build the wall. We need the wall.

Claim Checker - Knowledge Bases

Consulting the knowledge bases produced the following results:

Truth Rating Indeterminable

Question Asked

What is all of the problems-- the single biggest problem?

Response Received

The single biggest problem in communication is the illusion that it has ...

Claim Matcher

We found the following claims which have been professionally fact-checked. Check them out!

Truth Rating True

Claim

"Heroin ... pours across our southern borders."

Speaker Donald Trump

URL poltiffact

Truth Rating True

Claim Checker - Search Engine

We found the following information after processing some search engine results:

All of the problems -- the single biggest problem is heroin that pours across our southern border. It's just pouring and destroying their youth.

Similarity Rating 0.8320502943378437

URL source

"I was up in New Hampshire the other day." Trump said in the debate. **The biggest**

GENIUS

FILTER BY All Annotations

The biggest complaint they have -- it's with all of the problems going on in the world, many of the problems caused by Hillary Clinton and by Barack Obama

JosueCaraballo

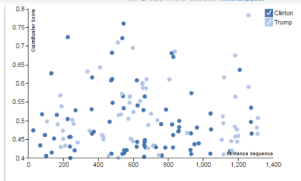
Is that true? What is the source?

report abuse

Like Upvote

Facebook Twitter

Add a comment



[?]

End-to-end system: FullFact.org [?]

LiveTrendsTools

BBC NewsBBC OneBBC TwoITVChannel 4PMQsRecExperimental ModeLive SubtitlesFull Fact

15:34:57Economists have said GDP is rising.

15:34:57Hospitals are struggling to provide essential gear, pi is the government gear, pi is the government cutting the number of beds in the National Health Service?

15:34:57-- beds. . The -- beds...

15:34:57The average length of time for staying in the hospital has virtually Haft since the year 2000.

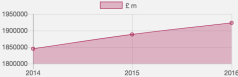
15:34:57We can look at the Labour Party 's record on this.

15:34:58In the last six years of the last Labour government, 25,000 hospital beds were cut.

Robo matches

Found automatic match

From 2015 to 2016, gdp rose. It went from 1888737.0 to 1922626.0, a percentage change of 1.8%.



Year	GDP (£ m)
2014	1888737.0
2015	1922626.0
2016	1922626.0

See the source data

Exact claim matches

The government is cutting the number of hospital beds.

This is right, the number of overnight beds in the English NHS has fallen.

See on the site | Trends | Tweet This

Was this the right claim to check? | | | Admin

Exact claim matches

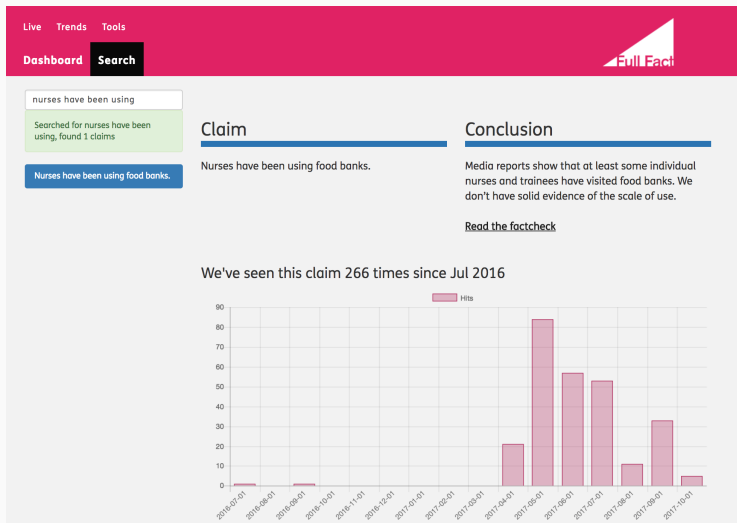
The average length of time for staying in hospital has virtually halved since the year 2000.

It hasn't quite halved but it has fallen. The average length of stay overnight was eight days in 2000 and is about five days now.

See on the site | Trends | Tweet This

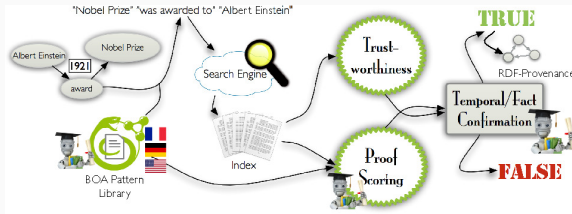
Was this the right claim to check? | | | Admin

End-to-end system: FullFact.org (continued)



[?]

End-to-end system: DeFacto [?]



- ▶ The system takes as input an **RDF Triple**, or a sentence that can be translated into one.
- ▶ Returns a set of **pages**, or **excerpts** thereof, w/ **source trustworthiness** (relying on PageRank, and page authority on a given topic)
- ▶ **Confidence score** computed based on the **number of proofs** found and **source trustworthiness**.
- ▶ Try to match the **triple** against the Linked Open Dataset
- ▶ The search for matches is done by **verbalizing** the input **RDF triples** and relying on **search engines**.

End-to-end systems: DeFacto (continued)

Charlie Sheen spouse Brooke Mueller

90.09% overall DeFacto score, fact holds for the year 2008 - 2011
263 websites containing the fact.

open proofs

spouse

74

Examples

Ahna O'Reilly, spouse, James Franco

Alexandra Christmann, spouse, Ben Kingsley

Alexis Valdés, spouse, Paulina Gálvez

Andrew Pruett, spouse, Abigail Spencer

Anna Torv, spouse, Mark Valley

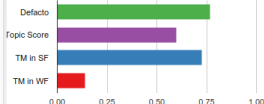
Blake Lively, spouse, Penn Dayton Badgley

Brian McFadden, spouse, Delta Goodrem

Brittany Murphy, spouse, Simon Monjack

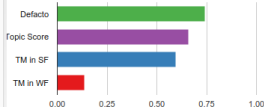
Carmine Giovinazzo, spouse, Vanessa Marcil

TMZ Live: Charlie Sheen -- Brooke Mueller's Blocking Our Sons ...



1. Z Live: **Charlie Sheen -- Brooke Mueller's Blocking Our Sons' Care** TMZ
2. **Charlie Sheen -- Brooke Mueller's Blocking Our Sons' Care** TMZ Live 11.

Charlie Sheen slams ex-wife Brooke Mueller on eve of her ...

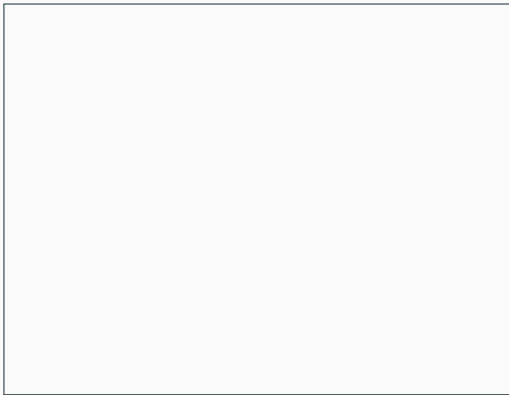


1. **Charlie Sheen slams ex-wife Brooke Mueller** on eve of her first unsu

Charlie Sheen's ex-wife Brooke Mueller completes ...



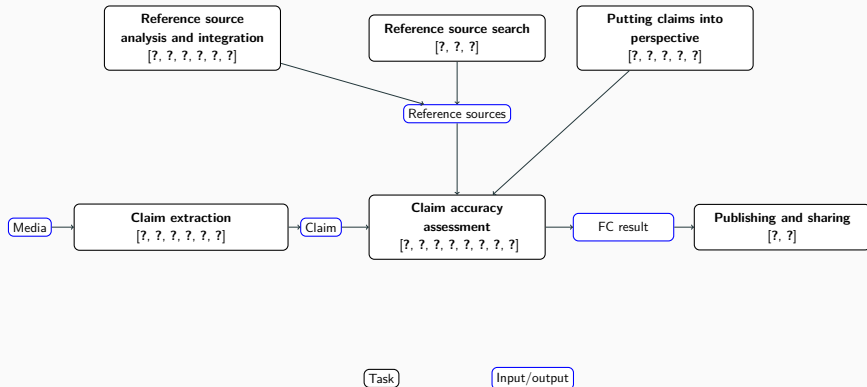
Online fact checking: Truthteller (2013)



Source: truthteller.washingtonpost.com
(now discontinued)

- ▶ **Now defunct**
- ▶ **Task:** Given a **video** of a discourse/debate, **identify claims** and link them to a **trusted source** of fact-checked claims (FactCheck.org).
- ▶ **Approach:** **Speech recognition** and basic **similarity** metrics between texts (from videos and from the trusted database).

Overview



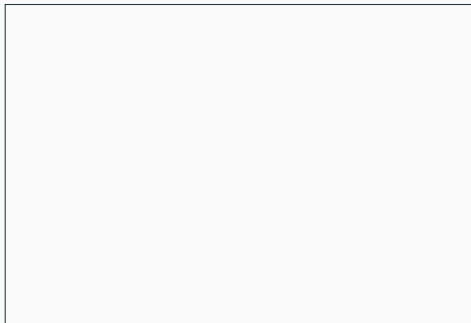
Perspectives

Agreed-upon notions on **event**, **issue**, **claim**, **context** and **stance** would help

- ▶ **validate** new approaches,
- ▶ **evaluate** their coverage and efficiency,
- ▶ **compare** their capabilities.

Quality control

- ▶ Facebook discontinued the “Disputed Stories” experiment, following complaint over quality and potential bias¹⁵



- ▶ Finer-grained check-worthiness recognition.
Current systems rate on a scale from factual to opinionated. It would be useful to rate how context-dependent a factual claim is. E.g., “This city’s taxes have gone up 20% since the last elections” cannot be checked without context [?].

¹⁵newsroom.fb.com/news/2017/12/news-feed-fyi-updates-in-our-fight-against-misinformation/

Transparency, interpretability, accountability

- ▶ **Transparency** is technically **easy**, but usually **not enough**

Example (Fake news detection)

Publishing the machine learning **model** for a **fake news detection** system goes in the right direction, but such models are **hardly interpretable**.

- ▶ **Interpretability** is **harder** to achieve and typically requires foundations [?, ?].

Example (Expert systems)

Expert systems use to have an “explain facility”. We probably need it back!

- ▶ **Accountability** concerns **ownership** of statements, i.e., who-said-what vs. who-reported-way. The vast literature on **provenance** likely has a role to play.

Collaboration

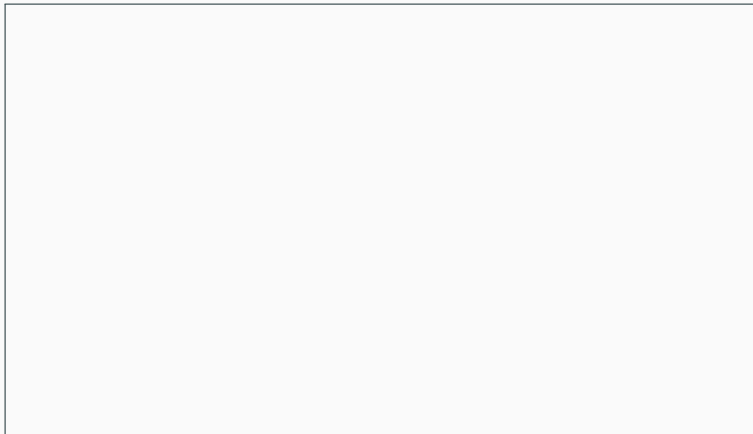
- ▶ From **exchanging** trusted data or previous fact checks, to **coordinated work** to face difficult investigations
- ▶ When fact checkers are of different sensibilities, fact checking becomes less partisan and credibility improves.
- ▶ Collaboration empowered by content management tools is a strong trend in journalism, promoted by organizations such as the ICIJ¹⁶
- ▶ CrossCheck¹⁷ is a premier example of this trend.

¹⁶International Consortium of Investigative Journalists, behind the Panama Papers and other such high-profile international investigations.

¹⁷<https://crosscheck.firstdraftnews.com/>

Collaboration (continued)

CJ Workbench: an online data curation and sharing platform for journalists.



Source: cjworkbench.org

- ▶ Beyond “ClaimReview” more standards are needed to cover **fact checking protocols, tools and functions**.
- ▶ A common and open framework for **naming issues and events**, and describing their interaction
- ▶ A common framework for managing **time** in Web data

Pluridisciplinarity

- ▶ **Social** and **cognitive sciences** useful to help devise psychologically effective fact-checking tools
- ▶ A recent whitepaper makes **recommendations** toward making fact checking more convincing, making it reach a larger audience, and avoiding viral misinformation¹⁸
- ▶ Interactions between computer scientists and journalists have been extremely fruitful for both sides [?]

Example (Computational lead finding)

An analysis of the way Wisconsin voting districts are drawn¹⁹, highlighting the (very) low probability that they may result from an “honest” design.

In the article's words, “it's math versus math, with democracy at stake”.

¹⁸americanpressinstitute.org/publications/reports/white-papers/future-of-fact-checking/

¹⁹nytimes.com/2017/10/06/opinion/sunday/computers-gerrymandering-wisconsin.html

Focus on issues over claims

- ▶ Most newsworthy questions are usually broader than just a claim. E.g., a misleading statement about criminal activity of refugees in the countries receiving them participates to a larger discussion about immigration, and the way different political parties argue it should be handled...
- ▶ One of the main points discussed in recent report by the American Press Institute^{??}.

Education

- ▶ **Data literacy**, envisioned as a set of **math** and **statistic skills**, through dedicated education modules at all levels.
- ▶ Some news outlets, e.g. France24's **The Observer**, have dedicated content on **critical thinking** and **news verification**.
- ▶ The **Google News Initiative** and **CrossCheck FR** now organize fact checking classes.
- ▶ Computer literacy is gaining ground in school curricula²⁰
Understanding the way media and communication works gives further tools to discern manipulation, statistic or otherwise.

²⁰See e.g. the course “Calling Bullshit: Data Reasoning in a Digital World” created at U. Washington, <http://callingbullshit.org/syllabus.html>

- ▶ Timely, sharp and balanced results.
- ▶ Avoiding frontal attack on one's convictions and beliefs.
- ▶ Choice of the best media for fact checking to reach each audience group.
- ▶ Engage and entertain the audience
 - ▶ Fact checking success is (also) judged by the audience it can gather and retain.

Data modeling and storage

Journalism is one of the **last industries to adopt digital tools**.

Many industries have successfully carried this transition, however for journalists, it is complicated by:

- ▶ Historical focus on **text** (not structured records)
- ▶ Strong focus on **creativity** and **speed** over **procedure, discipline, long-run**
- ▶ Lack of a **single application domain** (across the newroom); doable for specialized journalists or teams
- ▶ Limited financial means, with some notable exceptions (Ouest France)
- ▶ In many newrooms, **there is no long-term persistent content management plan** beyond archiving own articles
- ▶ A start: **reference databases**, e.g., of sports teams, precincts, public figures, companies...

What kinds of data journalists need to use?

- ▶ Whatever they can get their hands on
- ▶ Popular formats: PDF, JSON, CSV, XLS etc.
- ▶ Need automatic data types detection

All these data types need back-up mechanisms, e.g. (cloud-storage), CMS functionalities...

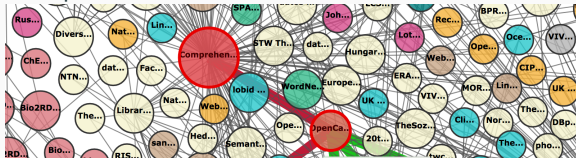
Data matching, linking and integration

Data comes in **heterogeneous data models, schemas**

Data sources **partially overlap** (or have similar topics) but have been produced **in isolation**.

This is the perfect setting in need of:

1. **Entity recognition**: identifying in text, mentions of a known structured entity
 - ▶ Link incoming text article to the entities it features, as they are described in the reference database
2. **Entity linking**: recognizing when two structured objects are the same
 - ▶ Well-known problem in the Web context



- ▶ Particular twist: strong bonus for **trusted data**

Natural Language Processing

- ▶ Basic NLP functionalities are used in some newsrooms but hardly exploited in workflows
 - ▶ **Named entity recognition** (person, location, organization names)
 - ▶ **Smart search**
 - ▶ **Voice recognition**

TITLE Prime Minister PERSON Justin Trudeau will interrupt his P9D DURATION nine-day foreign trip to return to CITY Ottawa

to try to end the deadlock over the ORGANIZATION Trans Mountain pipeline expansion .

Source: Stanford CoreNLP output

Natural Language Processing (continued)

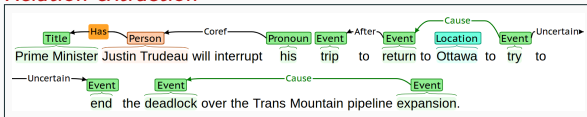
- ▶ Advanced approaches are not ready for production use by non-specialists, and require important human annotation effort:
 - ▶ Domain-specific **entity extraction**

LOAN AGREEMENT

This **LOAN AGREEMENT**, dated as of **November 17, 2014** (this “**Agreement**”), is made by and among **Auxilium Pharmaceuticals, Inc.**, a corporation incorporated under the laws of the State of **Delaware** (“**U.S. Borrower**”), **Auxilium UK LTD**, a private company limited by shares registered in **England and Wales** (“**UK Borrower**”) and, collectively with the **U.S. Borrower**, the “**Borrowers**”) and **Endo Pharmaceuticals Inc.**, a corporation incorporated under the laws of the State of **Delaware** (“**Lender**”).

Source: [?]

- ▶ **Relation extraction**



- ▶ **Knowledge discovery**

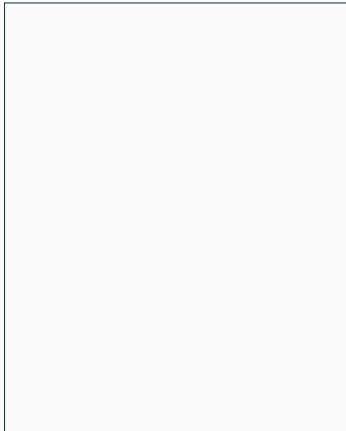
Natural Language Processing (continued)

- ▶ Bottlenecks currently tackled by NLP researchers:
 - ▶ **Data quality**: how to perform a good extraction from noisy or sparse data
 - ▶ **Data heterogeneity**: how to deal with knowledge distributed over structured, semi-structured and unstructured datasets.
 - ▶ **Supervision**: current effective approaches requires an important amount of human-annotated data. Reducing the need for human supervision is critical (distant supervision, active learning, domain adaptation, transfer learning, etc.)
 - ▶ **Reasoning** and **inference** is still limited.
 - ▶ **Interpretability** is a key challenge.
 - ▶ Industrial grade systems are still not the rule.
- ▶ Efficient NLP systems for fact checking will have to be **crosslingual**!

Time management

Almost **everything** is time-dependent

- ▶ Facts, beliefs and data evolves in time and have a limited period of validity.
- ▶ Events have start and end points.
- ▶ Fact check results become outdated, also!



The time dimension can be the news!

The time when someone does, say or learns something can make the difference between

- ▶ A willful **lie** or ignorance
- ▶ Lawful or **criminal behavior**, e.g., insider trading, lying to investigators

Example (Comey vs. Trump)

“you have to understand the chronology. The underlying question is whether Trump’s firing of Comey constituted obstruction of justice, which has a great deal to do with Flynn”²¹

Follows a chronology on **11 dates** and the conclusion:

“If we accept Comey’s account [...], then Trump asked Comey to drop the investigation of Flynn **after** members of his staff knew he had lied to the VP about it, and might even have had reason to believe he had lied to the FBI as well”.

²¹<https://www.washingtonpost.com/blogs/plum-line/wp/2018/04/20/heres-another-telling-revelation-in-the-comey-memos>

Time management (continued)

Tracing data and its evolution for accuracy, transparency, reproducibility.

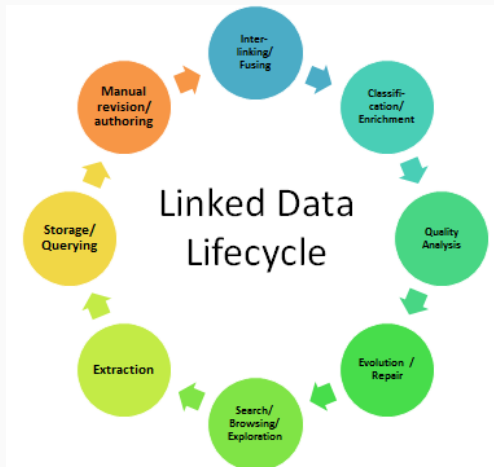
A FCMS should **record and permanently store time information** such as

- ▶ **Data creation** time stamp
- ▶ **Acquisition** times
- ▶ **Statement date**
- ▶ **Version** management
- ▶ In text, temporal expressions and their relations with events

Yet, despite a W3C recommendation (**OWL-Time**), there is no widely used standard for **representing time** in **Web data**.

Data quality management

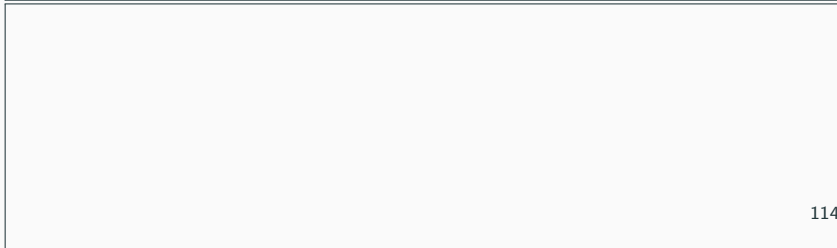
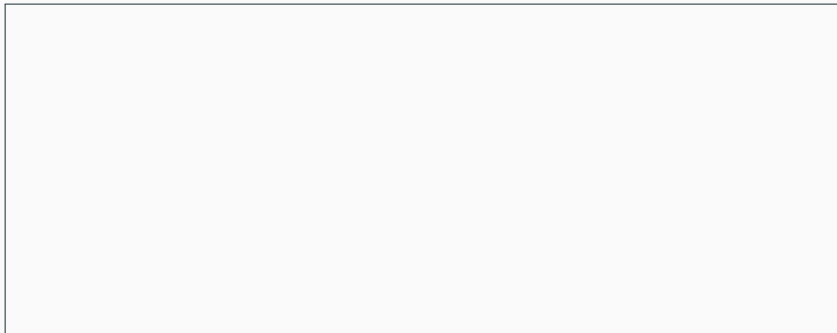
Applying **data life cycle** management tools to reference sources and fact checks.



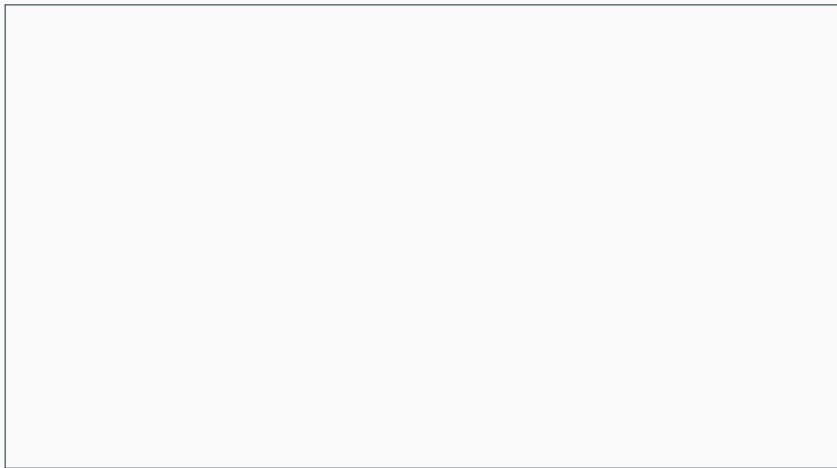
[?]

Enlisting experts

- ▶ Providing tools for experts to validate complex claims



Enlisting experts (continued)



Source: climatefeedback.org/

Support for reproducibility

Enabling to “replay” fact checking effort and get the same results.

- ▶ Fact checking can be seen as a **scientific** or **forensic** work
→ Reproducibility is needed
- ▶ This means:
 - ▶ **Defining** and **structuring** the fact-checking process, inputs and outputs
 - ▶ **Keeping trace** of manual fact-checking processes
 - ▶ Building multilingual **benchmarks** more complex than binary fake-news benchmarks
- ▶ Sharing reproducible results can help:
 - ▶ Strengthening a scientific community and accelerating the research.
 - ▶ Preserving (or regaining) citizens' trust.
 - ▶ Important for fact-checking in general, not only for automated solutions.

**Thank you for your attention.
Questions?**

Computational fact-checking

Problems, state of the art, and perspectives

Julien Leblay ¹ Ioana Manolescu ² Xavier Tannier ³

The Web Conference - April 2018

¹Artificial Intelligence Research Center – AIST, Japan

²Inria, LIX – CNRS and École Polytechnique, France

³Sorbonne Université – Inserm – LIMICS, France