Fake News Research: Theories, Detection Strategies, and Open Problems

Introduction

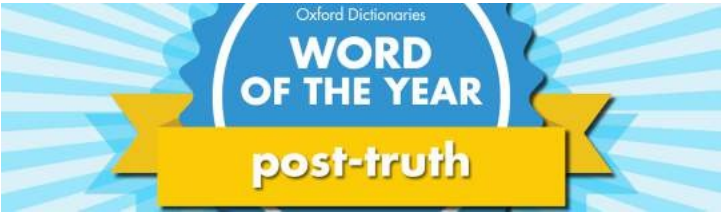
- Research Background
- What is Fake News?
- Related Concepts
- Fundamental Theories

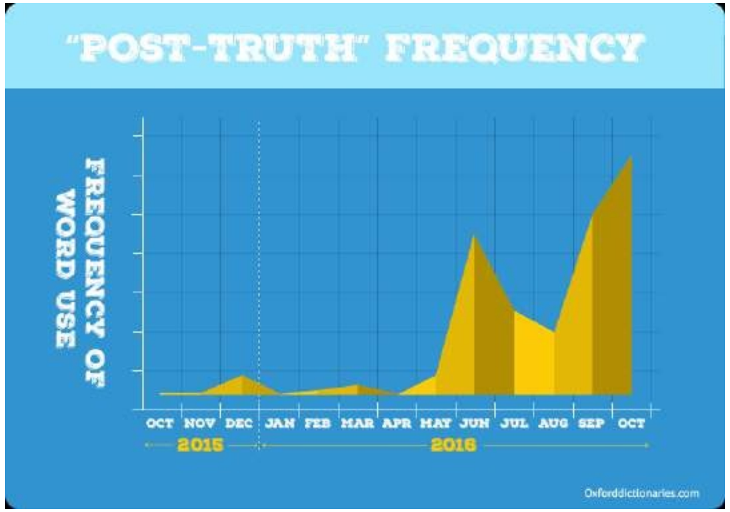
Research Background

Why Study Fake News?

Fake news is now viewed as one of the greatest threats to **democracy**, **justice**, **public trust**, **freedom of expression**, **journalism** and **economy**.

- **Political** Aspects: May have had an impact on
  - “Brexit” referendum
  - 2016 U.S. presidential election
    - # Shares, reactions, and comments on Facebook.[1]
    - 8,711,000 for top 20 frequently-discussed **FAKE** election stories.
    - 7,367,000 for top 20 frequently-discussed **TRUE** election stories.
- Oxford Dictionaries international word of the year 2016:
  - **Post-Truth**: “Relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief.”



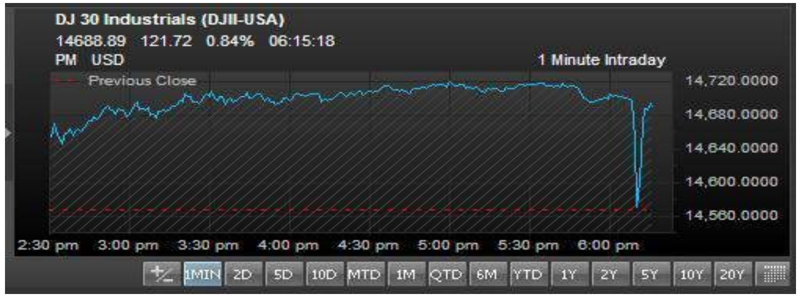


[1]C. Silverman. This analysis shows how viral fake election news stories outperformed real news on Facebook. BuzzFeed News, 2016.

Research Background

Why Study Fake News?

- **Economic** Aspects:
  - "Barack Obama was injured in an explosion" wiped out $130 billion in stock value.[1]
  - Dozens of "well-known" teenagers in Veles, Macedonia[2]
    - Penny-per-click advertising
    - During U.S. 2016 presidential Elections
    - Earning at least $60,000 in six months
    - Far outstripping their parents' income
    - Average annual wage in town: $4,800



[1] K. Rapoza. Can 'fake news' impact the stock market? 2017.
[2] S. Subramanian, Inside the Macedonnian Fake News Complex https://www.wired.com/2017/02/veles-macedonia-fake-news/

Research Background

Why Study Fake News?

- **Social/Psychological** Aspects:
  - Humans have been proven to be irrational/vulnerable when differentiating between truth/false news
    - Typical accuracy in the range of 55-58%
  - For fake news, it is relatively easier to obtain public trust
  - **Validity Effect:** individuals tend to trust fake news after repeated exposures
  - **Confirmation Bias:** individuals tend to believe fake news when it confirms their pre-existing knowledge
  - **Peer Pressure/Bandwagon Effect**

Research Background

Why is Fake News attracting more public attention recently?

- Fake news can now be created and *published faster* and *cheaper*
- The rise of Social Media and its popularity also plays an important role
  - As of Aug. 2017, 67% of Americans *get* their news from social media.[3]
- Social media accelerates *dissemination* of fake news.
  - It breaks the physical distance barrier among individuals.
  - It provides rich platforms to share, forward, vote, and review to encourage users to participate and discuss online news.
- Social media accelerates *evolution* of fake news.
  - **Echo chamber effect**: biased information can be amplified and reinforced within the social media.[4]
  - **Echo Chamber:** a situation in which beliefs are amplified or reinforced by communication and repetition inside a closed system

[3] http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/
[4] K. Jamieson and J. Cappella. Echo Chamber: Rush Limbaugh and the Conservative Media Establishment. Oxford University Press, 2008.

What Is Fake News?

Fake News & Related Concepts

Definition of fake news

*Fake news is **intentionally** and verifiably **false** news published by a **news** outlet.*

- *Intention*: Bad
- *Authenticity*: False
- *News or not?* News

A more broad definition:

- *Fake news is false news*

	Authenticity	Intention	News?
Fake news	False	Bad	Yes
False news	False	Unknown	Yes
Satire news	Unknown	Not bad	Yes
Disinformation	False	Bad	Unknown
Misinformation	False	Unknown	Unknown
Rumor	Unknown	Unknown	Unknown

For example, disinformation is false information [news or non-news] with a bad intention aiming to mislead the public.

Fake News & Related Concepts

Distinguishing fake news from other related concepts

Open Problems:

- How similar are writing styles or propagation patterns?
- Can we use the same detection strategies?
- Can we distinguish between them? E.g., fake news from satire news

(a) Style
(b) propagation

Fundamental Theories

Style-Based Fundamental Theories

Studying fake news from a style perspective, i..e, how it's written

	Term	Phenomenon
Style-based	*Undeutsch hypothesis*	A statement based on a factual experience differs in **content and quality** from that of fantasy
	Reality monitoring	Actual events are characterized by higher levels of **sensory-perceptual** information.
	Four-factor theory	Lies are expressed differently in terms of arousal, behavior control, **emotion**, and thinking from truth.

w/o really fake

behind the design of traditional lie

R. Zafarani, X. Zhou, K. Shu, H. Liu

Propagation-based Fundamental Theories

Studying fake news based on how it spreads

	Term	Phenomenon
Propagation - based	*Backfire effect*	Given evidence against their beliefs, individuals can reject it even more strongly
	Conservatism bias	The tendency to revise one's belief insufficiently when presented with new evidence.
	Semmelweis reflex	Individuals tend to reject new evidence as it contradicts with established norms and beliefs.

"Fake news is incorrect but hard to correct" [5]

It is difficult to correct users' perceptions after fake news has gained their trust.

Fake News Early Detection!

Providing a solid foundation for epidemic models

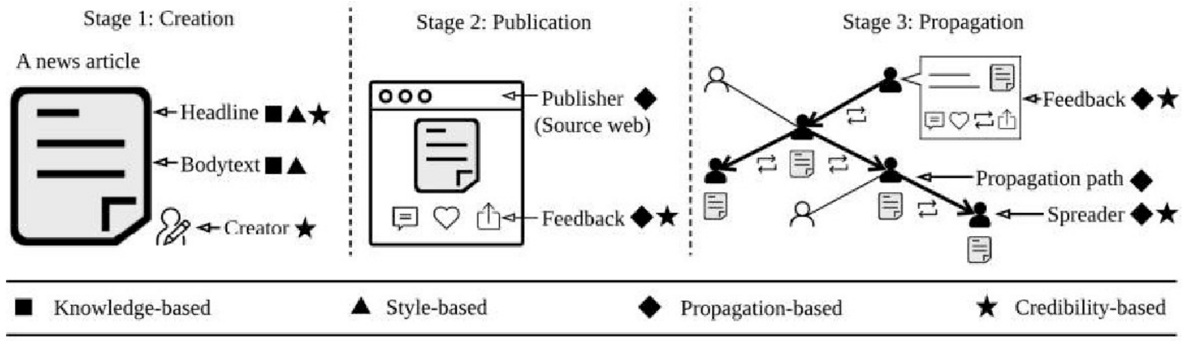
[5] A. Roets, et al. 'Fake news': Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions. Intelligence, 2017.

Fake News Detection

- Knowledge-based Fake News Detection
- Style-based Fake News Detection
- Propagation-based Fake News Detection
- Credibility-based Fake News Detection
- Fake News Datasets & Tools

Fake News Detection

- Knowledge-based Fake News Detection
- Style-based Fake News Detection
- Propagation-based Fake News Detection
- Credibility-based Fake News Detection
- Fake News Datasets & Tools



R. Zafarani, X. Zhou, K. Shu, H. Liu

Knowledge-based Fake News Detection

Overview

Knowledge-based fake news detection aims to assess **news authenticity** by comparing the **knowledge** extracted from to-be-verified **news content** with known facts (i.e., true knowledge).

It is also known as **fact-checking**.

- *Manual Fact-checking* – providing ground truth.
- *Automatic Fact-checking* – a better choice for scalability.

R. Zafarani, X. Zhou, K. Shu, H. Liu

Manual Fact-checking

Classification and comparison

	Expert-based manual fact-checking	**Crowd-sourced manual fact-checking**
Fact-checker(s)	One or several domain-expert(s)	A large population of regular individuals
Easy to manage?	Yes	No
Credibility	High	Comparatively low
Scalability	Poor	Comparatively high
Current resources (e.g., websites)	Rich	Comparatively poor

E.g., political bias and conflicting annotations of fact-checkers

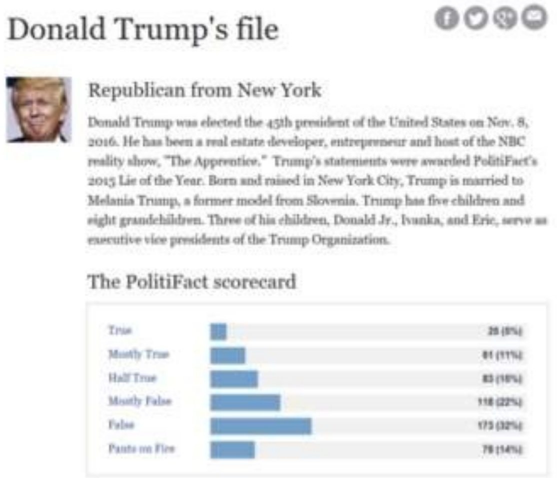
Expert-based Manual Fact-checking

Current resources

Multilabel classification

Binary classification

	Topics Covered	Content Analyzed	Assessment Labels
PolitiFact	American politics	Statements	True; Mostly true; Half true; Mostly false; False; Pants on fire
The Washington Post Fact Checker	American politics	Statements and claims	One pinocchio; Two pinocchio; Three pinocchio; Four pinocchio; The Geppetto checkmark; An upside-down Pinocchio; Verdict pending
FactCheck	American politics	TV ads, debates, speeches, interviews and news	True; No evidence; False
Snopes	Politics and other social and topical issues	News articles and videos	True; Mostly true; Mixture; Mostly false; False; Unproven; Outdated; Miscaptioned; Correct attribution; Misattributed; Scam; Legend
TruthOrFiction	Politics, religion, nature, aviation, food, medical, etc.	Email rumors	Truth; Fiction; etc.
FullFact	Economy, health, education, crime, immigration, law	Articles	Ambiguity (no clear labels)
HoaxSlayer	Ambiguity	Articles and messages	Hoaxes, scams, malware, bogus warning, fake news, misleading, true, humour, spams, etc.



across domains

Multi-modal

Knowledge-based Fake News Detection

Overview

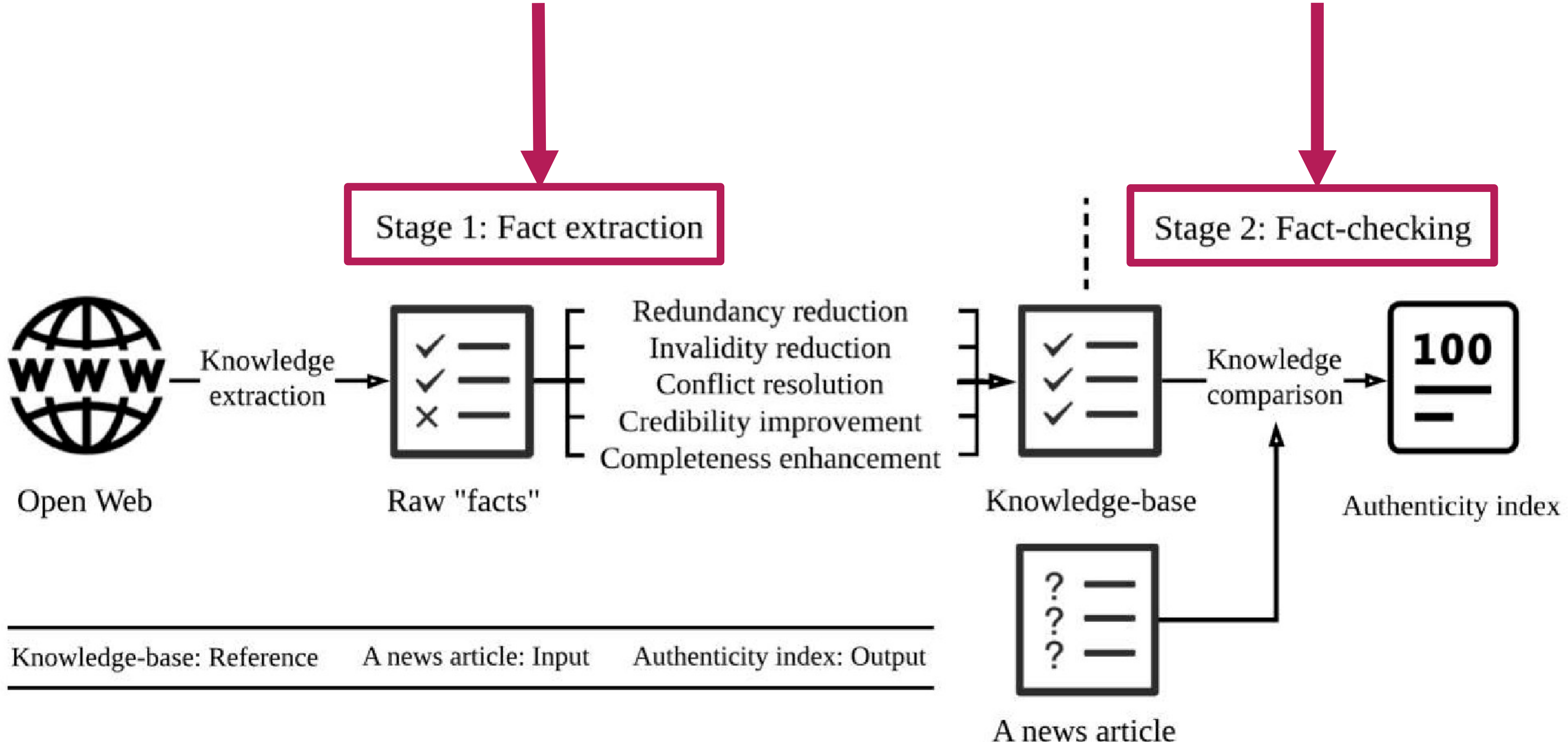
Knowledge-based fake news detection aims to assess **news authenticity** by comparing the **knowledge** extracted from to-be-verified **news content** with known facts (i.e., true knowledge).

It is also known as **fact-checking**.

- *Manual Fact-checking* – providing ground truth.
- *Automatic Fact-checking* – a better choice for scalability.

It aims to assess news authenticity by comparing the knowledge extracted from to-be-verified news content with known facts (i.e., true knowledge).

- How to represent "**knowledge**"?
- How to obtain **the known facts** (i.e., ground truth)?
- How to **compare** the knowledge extracted with known facts?



Automatic Fact-checking

Overview

Manual Automate

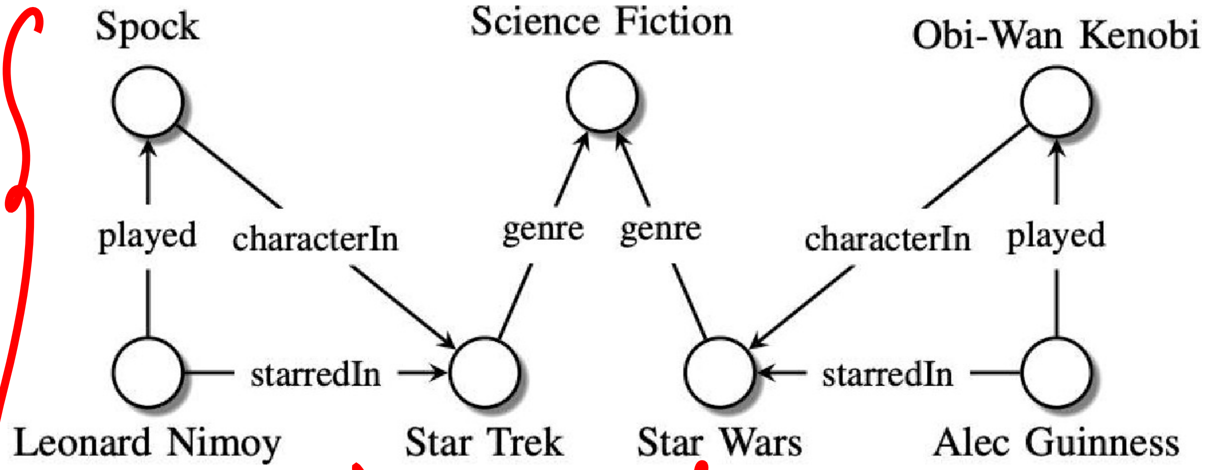
Relation

Knowledge Representation

Knowledge is represented as **a set of (Subject, Predicate, Object) (SPO) triples** extracted from the given information. For example,

"Leonard Nimoy was an actor who played the character Spock in the science-fiction movie Star Trek"

subject	*predicate*	*object*
(LeonardNimoy,	*profession,*	*Actor)*
(LeonardNimoy,	*starredIn,*	*StarTrek)*
(LeonardNimoy,	*played,*	*Spock)*
(Spock,	*characterIn,*	*StarTrek)*
(StarTrek,	*genre,*	*ScienceFiction)*



Knowledge Graph.

The illustration is from: M. Nickel, et al. A Review of Relational Machine Learning for Knowledge Graphs, 2016

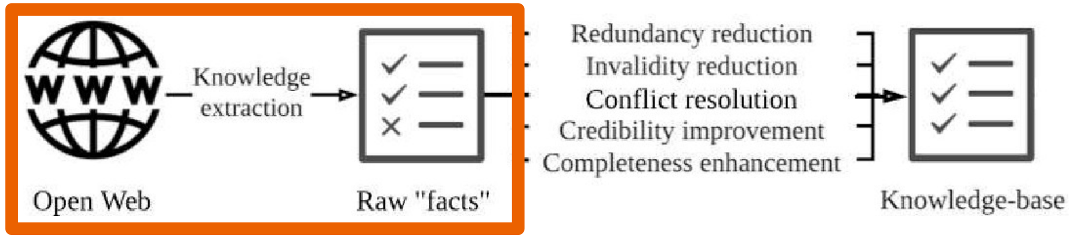
Stage 1. Fact Extraction

Constructing knowledge graph to obtain the known facts

Types of Web content that contain relational information and can be utilized for knowledge extraction by different extractors: **text, tabular data, structured pages** and **human annotations.**[6]

Source(s):

- Single-source knowledge extraction
  - Rely on one comparatively reliable source (e.g., Wiki)
  - Efficient ⬆,Knowledge completeness ⬇
- Open-source knowledge extraction
  - Fuse knowledge from distinct knowledge
  - Efficient ⬇,Knowledge completeness ⬆



[6]X. Dong, et al.. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. KDD'14

T1: **Entity Resolution (deduplication/record linkage)** to reduce redundancy

- To identify mentions that refer to the same real-world entity, e.g., *(DonaldJohnTrump,profession, President)* & *(DonaldTrump, profession, President)* should be a redundant pair.
- Current techniques are often distance- or dependence-based.
- Often expensive (requires pairwise distance) computation
- Blocking/Indexing can be used to deal with complexity

T2: **Time Recording** to remove outdated knowledge

- E.g., *(Britain, joinIn, EuropeanUnion)* has been outdated.
- Use Compound Value Type (CVT): facts having beginning and end dates
- Timeliness studies are limited

X → out of the KB/KG

T3: **Knowledge Fusion** to handle conflicts (often in open-source knowledge extraction)

- E.g., *(DonaldTrump, bornIn, NewYorkCity)* & *(DonaldTrump, bornIn, LosAngeles)* are a conflicting pair.
- Fix by having support values for facts (e.g., website credibility), or using ensemble methods
- Often correlated to (T4).

T4: **Credibility Evaluation** to improve the credibility of knowledge

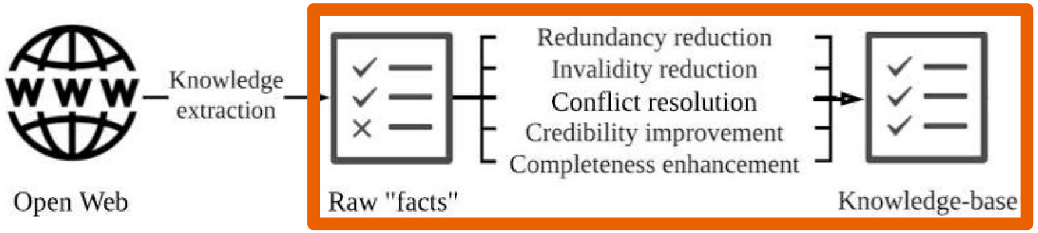
- E.g., The knowledge extracted from The Onion[7].
- Often focus on analyzing the source website(s).

[7] A https://www.theonion.com/

new Onion

Stage 1. Fact Extraction

Constructing knowledge graph to obtain the known facts



T5: ***Knowledge Inference/Link Prediction*** to infer new facts based on known ones

- Knowledge extracted from online resources, particularly, using a single source, are far from complete.

Relation machine learning

Latent Feature Models, e.g., **RESCAL**

Assume the existence of knowledge-base triples is conditionally independent given latent features and parameters

Graph Feature Models, e.g., **PRA**

Assume the existence of triples is conditionally independent given observed graph features and parameters

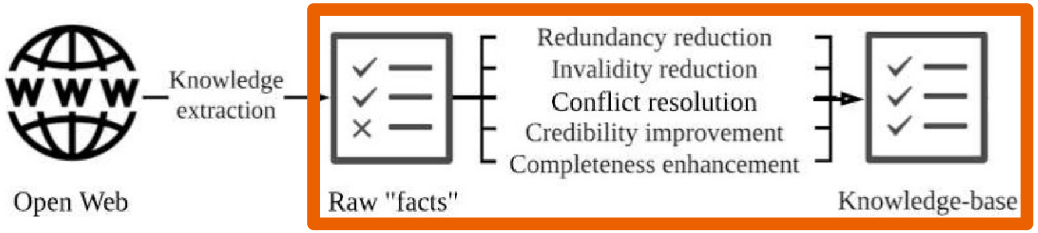
Markov Random Field (MRF) Models

Assume the existing triples have local interactions

M. Nickel, et al. A Review of Relational Machine Learning for Knowledge Graphs, Proceedings of the IEEE, 2016

Stage 1. Fact Extraction

Constructing knowledge graph to obtain the known facts



Stage 1. Fact Extraction

Existing Knowledge Graphs

Name
Knowledge Vault (KV)
DeepDive [32]
NELL [8]
PROSPERA [30]
YAGO2 [19]
Freebase [4]
Knowledge Graph (KG)

Table 1: Comparison of
Freebase and KG rely o
facts means with a prob

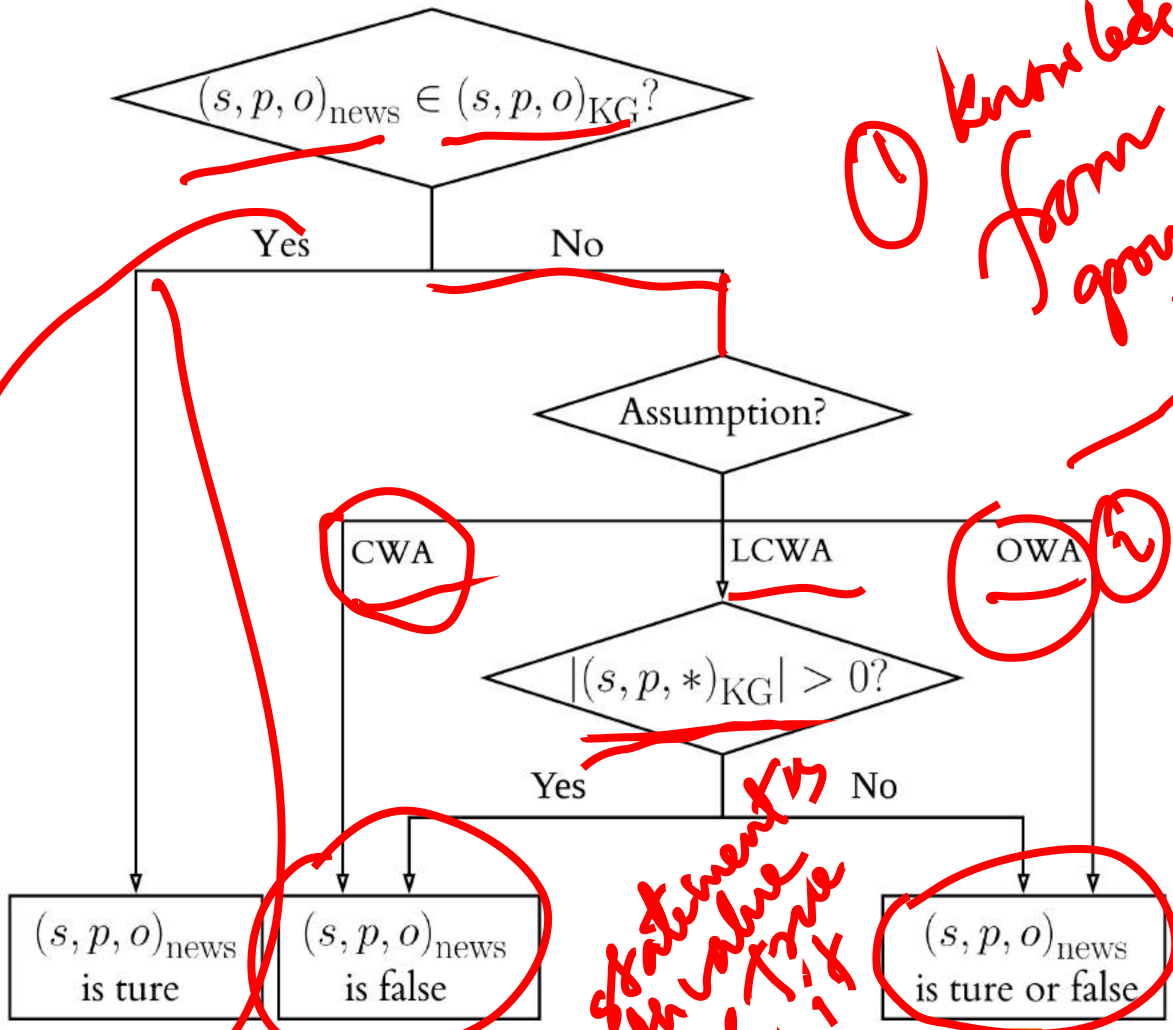
[a]Ce Zhang (U Wisconsin), private communication
[b]Bryan Kiesel (CMU), private communication
[c]Core facts, `http://www.mpi-inf.mpg.de/yago-naga/yago/downloads.html`
[d]This is the number of non-redundant base triples, excluding reverse predicates and "lazy" triples derived from flattening CVTs (complex value types).
[e]`http://insidesearch.blogspot.com/2012/12/get-smarter-answers-from-knowledge_4.html`

Open issues:

1. **Timeliness & Completeness of Knowledge Graphs**
2. **Domain-specific Knowledge Graphs for Fake News Detection**

Related tutorial: X. Ren, et al., Scalable Construction and Querying of Massive Knowledge Bases, WWW tutorial, 2018.

Source: X. Dong, et al.. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. KDD'14

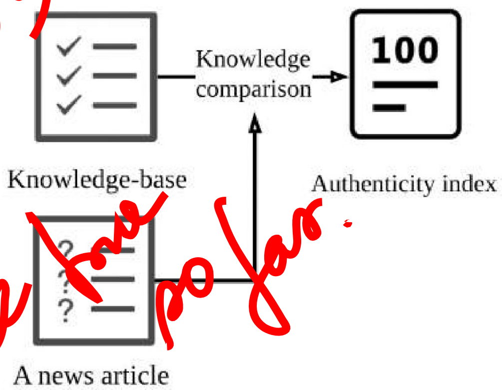


KG: Knowledge Graph
CWA: Closed-World Assumption
LCWA: Local Closed-World Assumption
OWA: Open-World Assumption

Knowledge Inference

Stage 2. Fact-checking

Comparing knowledge between news articles and knowledge graphs



① knowledge from the ground-truth KG

② knowledge from the input news article (s,p,o) if it is known to be true so far.

a statement's truth value can be true only if it is known to be true.

statement's truth value can be true even if not

Ascertain

u Barack Obama is of religion Muslim

wiki inforbox

Shortest path-based method:

By finding the **shortest path** between concept nodes under properly defined **semantic proximity** metrics on knowledge graphs

$v_i \rightarrow$ concept node

edge-relation

$k(v_i)$ degree

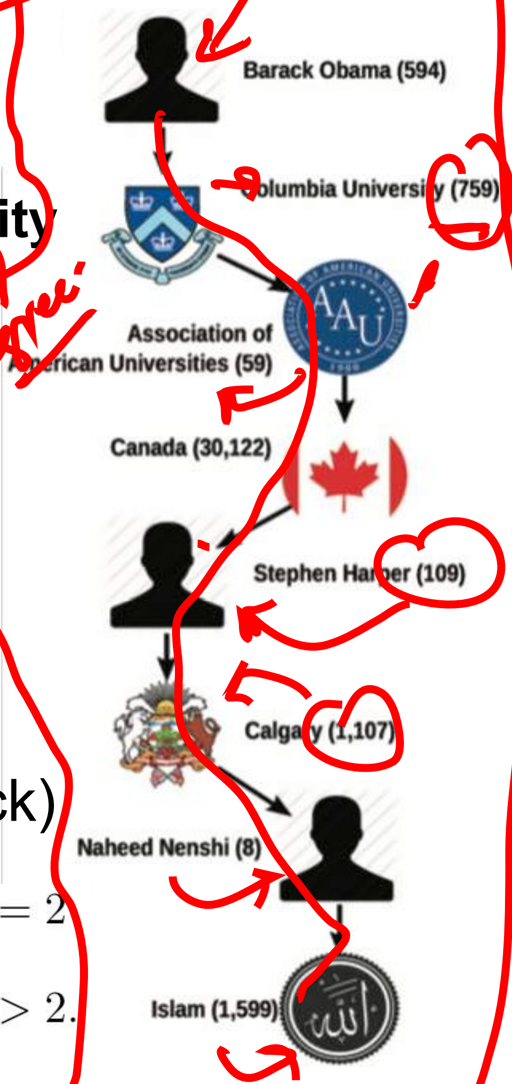
$$\tau(e) = \max \mathcal{W}(P_{s,o}).$$

$$\mathcal{W}(P_{s,o}) = \mathcal{W}(v_1 \ldots v_n) = \left[1 + \sum_{i=2}^{n-1} \log k(v_i)\right]^{-1}$$

An alternative formulation (widest bottleneck)

$$\mathcal{W}_u(P_{s,o}) = \mathcal{W}_u(v_1 \ldots v_n) = \begin{cases} 1 & n = 2 \\ \left[1 + \max_{i=2}^{n-1} \{\log k(v_i)\}\right]^{-1} & n > 2. \end{cases}$$

Atomic Bar



G. Ciampaglia, et al. Computational Fact Checking from Knowledge Networks, 2016

Stage 2. Fact-checking

Knowledge Inference for unknown SPO triples: Illustrated studies

FakeNewsTracker: A Tool for Fake News Collection, Detection, and Visualization

Kai Shu, Deepak Mahudeswaran, and Huan Liu

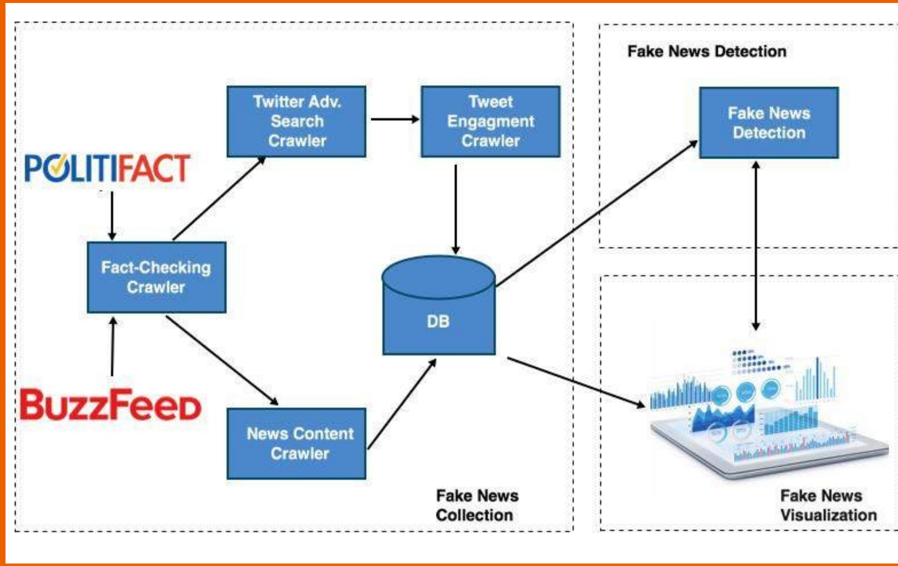
SBP 2018

SBP Disinformation Challenge Winner

http://blogtrackers.fulton.asu.edu:3000

An end-to-end framework for fake news collection, detection, and visualization

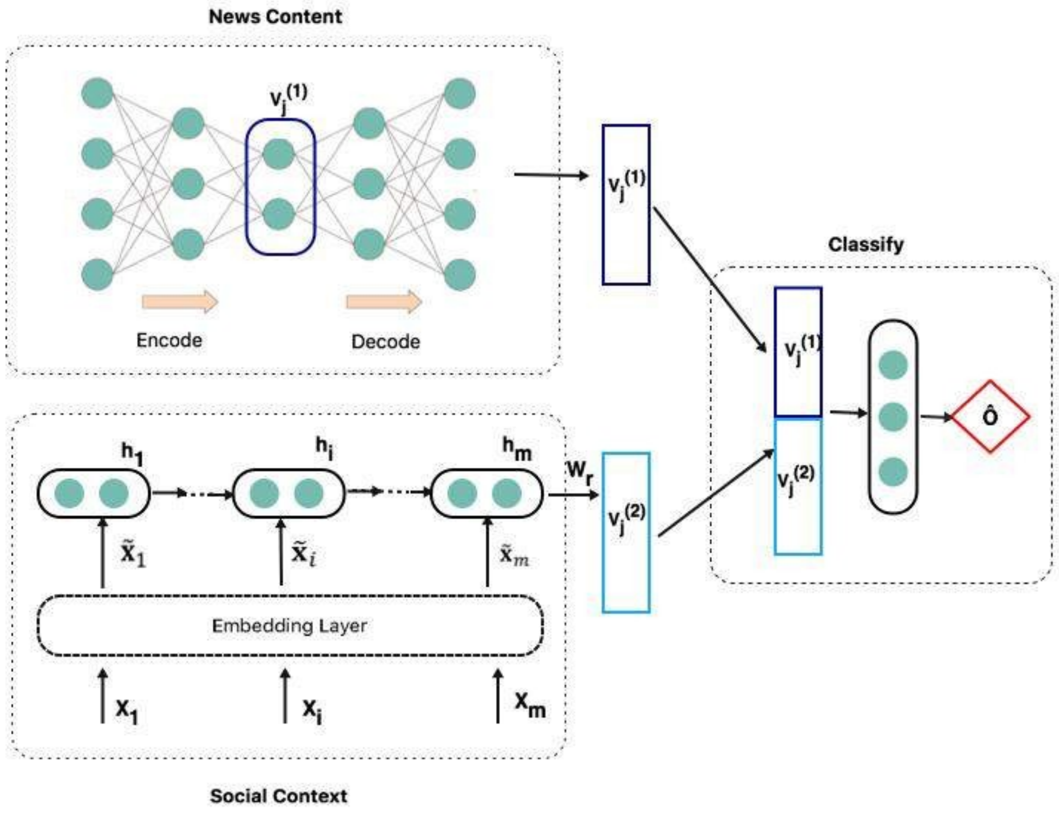
- **Data Collection:** collecting fake and real news articles from fact-checking websites and related social engagements from social media
- **Fake News Detection**: finding fake news with advanced machine learning methods, such as deep neural networks
- **Fake News Visualization:** visualization on data attributes and model performance



X. Zhou, R. Zafarani, K. Shu, H. Liu

Fake News Detection

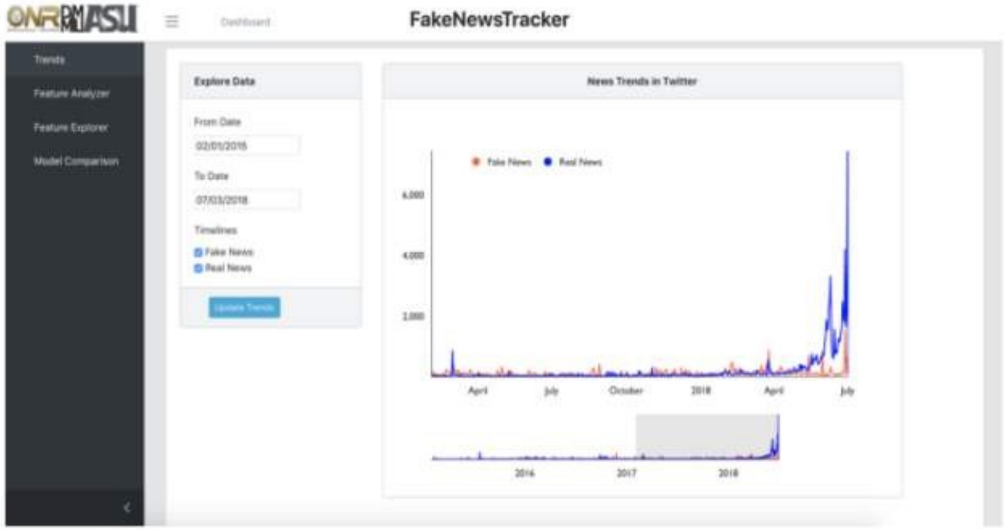
- Detect fake news with fusion of news content and social context
  - **News representation**: Represent news content using autoencoders
  - **Social engagement representation:** Represent social engagements using RNNs
  - **Social Article Fusion:** Combine both news and social engagement features to detect fake news



X. Zhou, R. Zafarani, K. Shu, H. Liu

Fake News Visualization

Trends on Twitter



Topics of Fake news vs Real News

Geolocation of Fake News vs Real News

Social Network on Users Spreading Fake/Real news

Potential Applications for FakeNewsNet

- **Fake News Detection**
  - News content, social context based
  - Early fake news detection
- **Fake News Evolution**
  - Temporal, Topic, Network, evolution
- **Fake News Mitigation**
  - Provenances, persuaders, clarifiers
  - Influence minimization, mitigation campaign
- **Malicious Account Detection**
  - Detecting bots that spread fake news