Tutorial @ AAAI 2018, New Orleans, LA, USA

Computational Solutions against Fake News: AI vs. DB Approaches

Naeemul Hassan
nhassan@olemiss.edu
University of Mississippi

Dongwon Lee
dongwon@psu.edu
Penn State University

Part 3. DB Approaches

Outline

- Challenges
- Fake News & Fact-checking
- Tracking Misinformation
- Crowdsourced Fact-checking
- Open Questions

Available Datasets

- ASU FakeNewsNet: https://github.com/KaiDMML/FakeNewsNet
- PHEME rumor: https://goo.gl/5JXAXt
- NewsVerification: https://goo.gl/jDZJpS
- BuzzFeedNews: https://goo.gl/TUnw2K
- LIAR: https://goo.gl/wkNYhv
- BS Detector: https://github.com/bs-detector/bs-detector
- FakeNewsSites: https://www.kaggle.com/mrisdal/fake-news/data

Challenges

- There are no agreed upon bench-mark datasets for the fake news detection problem. [Shu et al., 2017]

Challenges

- There are no agreed upon bench-mark datasets for the fake news detection problem. [Shu et al., 2017]

BS Detector[1]: This dataset is collected from a browser extension called BS detector developed for checking news veracity. It searches all links on a given web-page for references to unreliable sources by checking against a manually complied list of domains. The labels are the outputs of BS detector, rather than human annotators.

1. https://www.kaggle.com/mrisdal/fake-news

Challenges

- There are no agreed upon bench-mark datasets for the fake news detection problem. [Shu et al., 2017]

LIAR[1]: This dataset is collected from fact-checking website PolitiFact through its API. It includes 12,836 human-labeled short statements. The labels for news truthfulness are fine-grained multiple classes: pants-fire, false, barely-true, half-true, mostly true, and true.

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

Challenges

- There are no agreed upon bench-mark datasets for the fake news detection problem. [Shu et al., 2017]

CREDBANK[1]: This is a large scale crowdsourced dataset of approximately 60 million tweets. All the tweets are broken down to be related to over 1,000 news events, with each event assessed for credibility by 30 annotators from Amazon Mechanical Turk

1. http://compsocial.github.io/CREDBANK-data/

Comparison of Fake News Detection Datasets

Features / Dataset	News Content		Social Context		
	Linguistic	Visual	User	Post	Network
BuzzFeedNews	✓				
LIAR	✓				
BS Detector	✓				
CREDBANK	✓		✓	✓	✓

FakeNewsNet[1]

To address the disadvantages of existing fake news detection datasets, [Shu et al., 2017] have an ongoing project to develop a usable dataset for fake news detection on social media. This dataset includes all mentioned news content and social context features with reliable ground truth fake news labels.

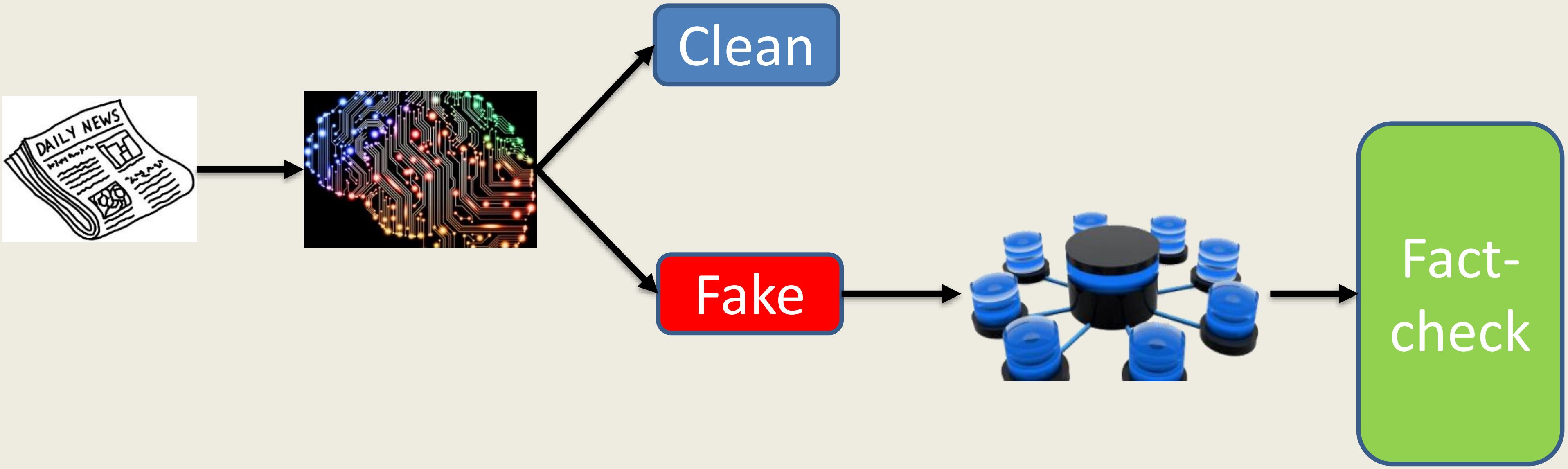
- Source
- Headline
- Body
- Image

1. https://github.com/KaiDMML/FakeNewsNet

Fake News &
Fact-checking

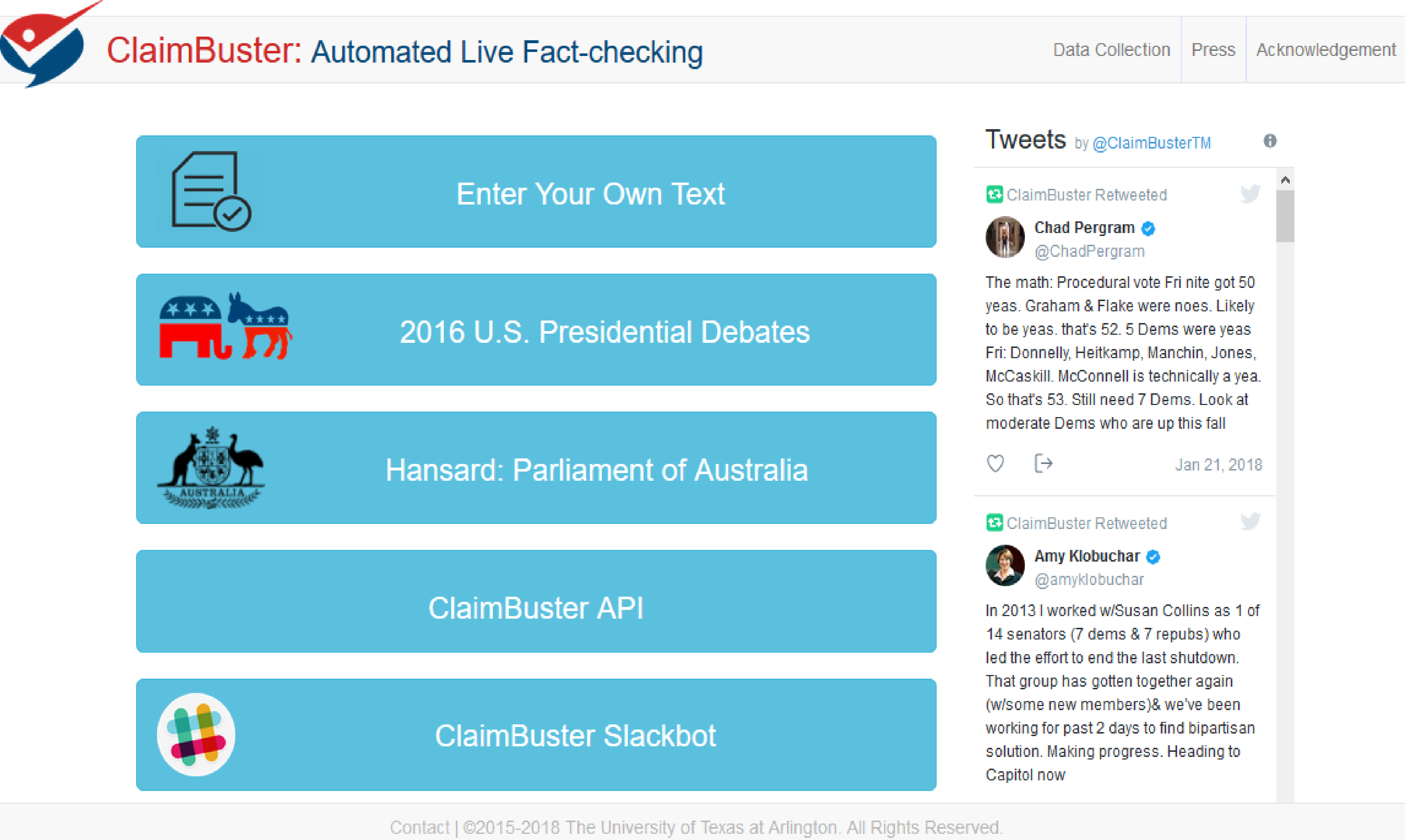
Fake News & Fact-checking

- Machine learning is good at finding patterns. AI can raise the flag. But to check it’s veracity, we need to check against a knowledge base.



Challenges

- Linking Multiple Data sources.



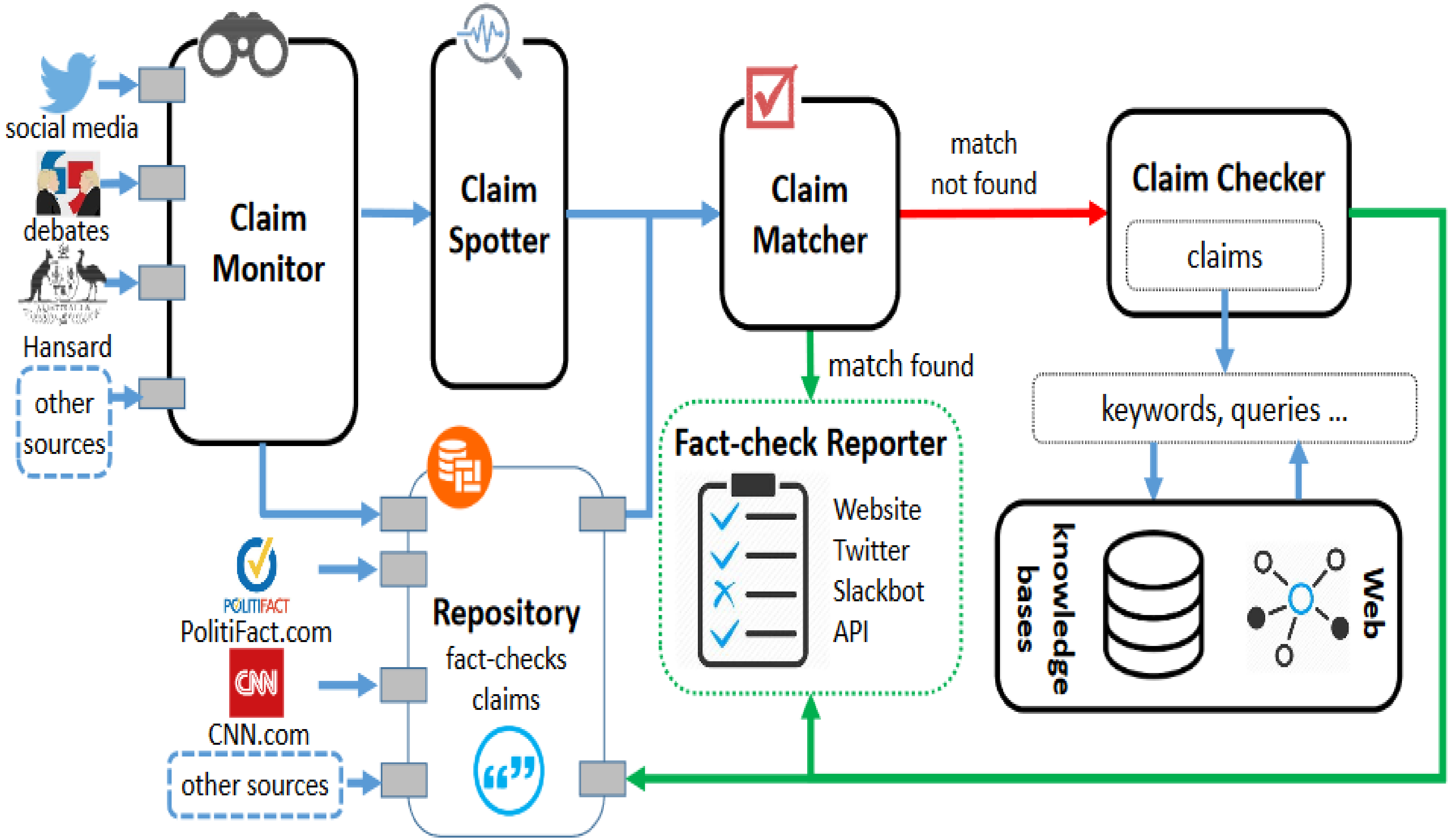
http://idir.uta.edu/claimbuster

[Hassan et al., 2017a, Hassan et al., 2017b]

ClaimBuster[1,2,3] Demo Time

1. Go to http://idir.uta.edu/claimbuster
2. Claim Spotter
    1. Any text
    2. 2016 U.S. Presidential Election
    3. Social media [Twitter]
    4. Hansard
3. Fact-Checker
    1. Knowledge base [Wolfram Alpha, Google, Freebase]
    2. Search Engines
    3. Fact-checking Platforms

ClaimBuster System Architecture



ClaimBuster: Fact to Question

Fact: The capital of Louisiana is New Orleans.

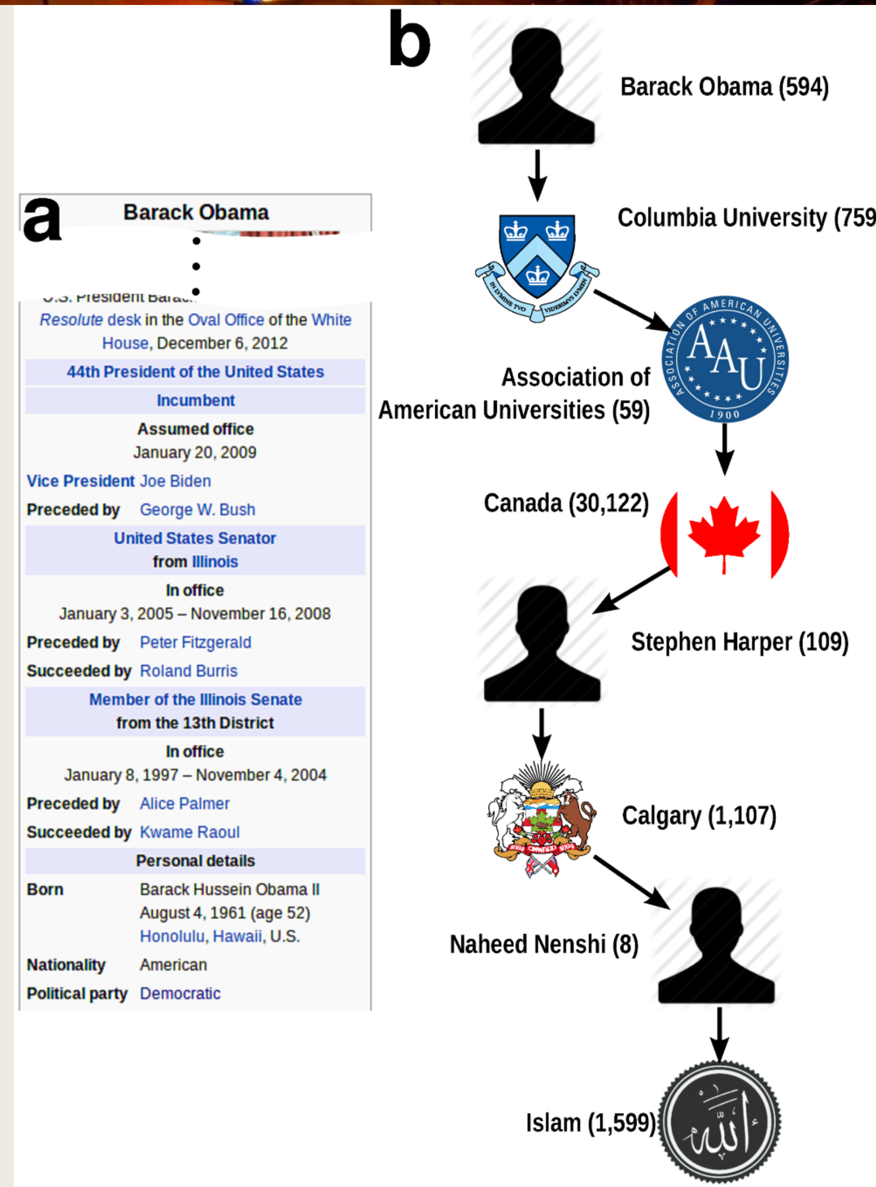
Question:

- What is the capital of Louisiana?
- What is New Orleans?
- Is the capital of Louisiana New Orleans?

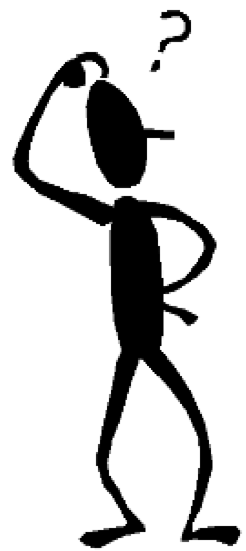
Knowledge Network Approach

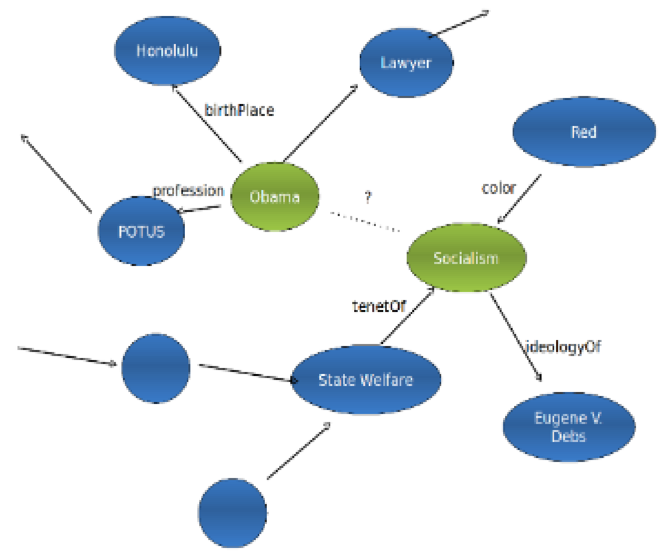
- Create a knowledge graph from Wikipedia.
- Veracity of a subject-predicate-object statement → finding a path between subject and object entities.

[Ciampaglia et al., 2015]



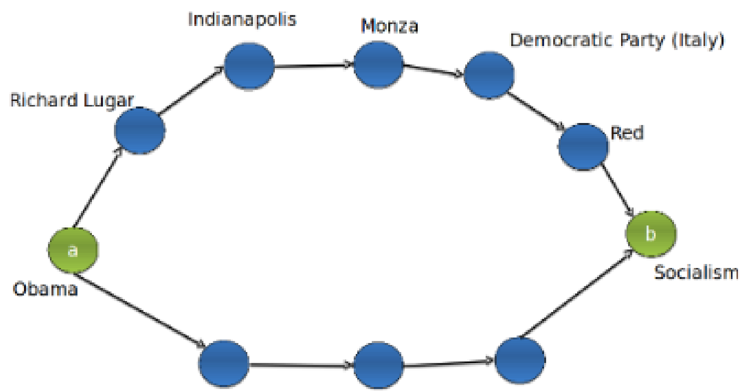
- Amount of "work" needed to resolve a novel statement
- On a network:
  how hard it is to find a path between two nodes
- Graph closure

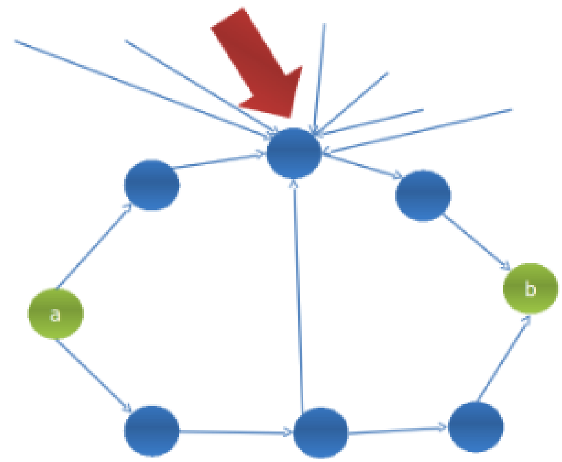




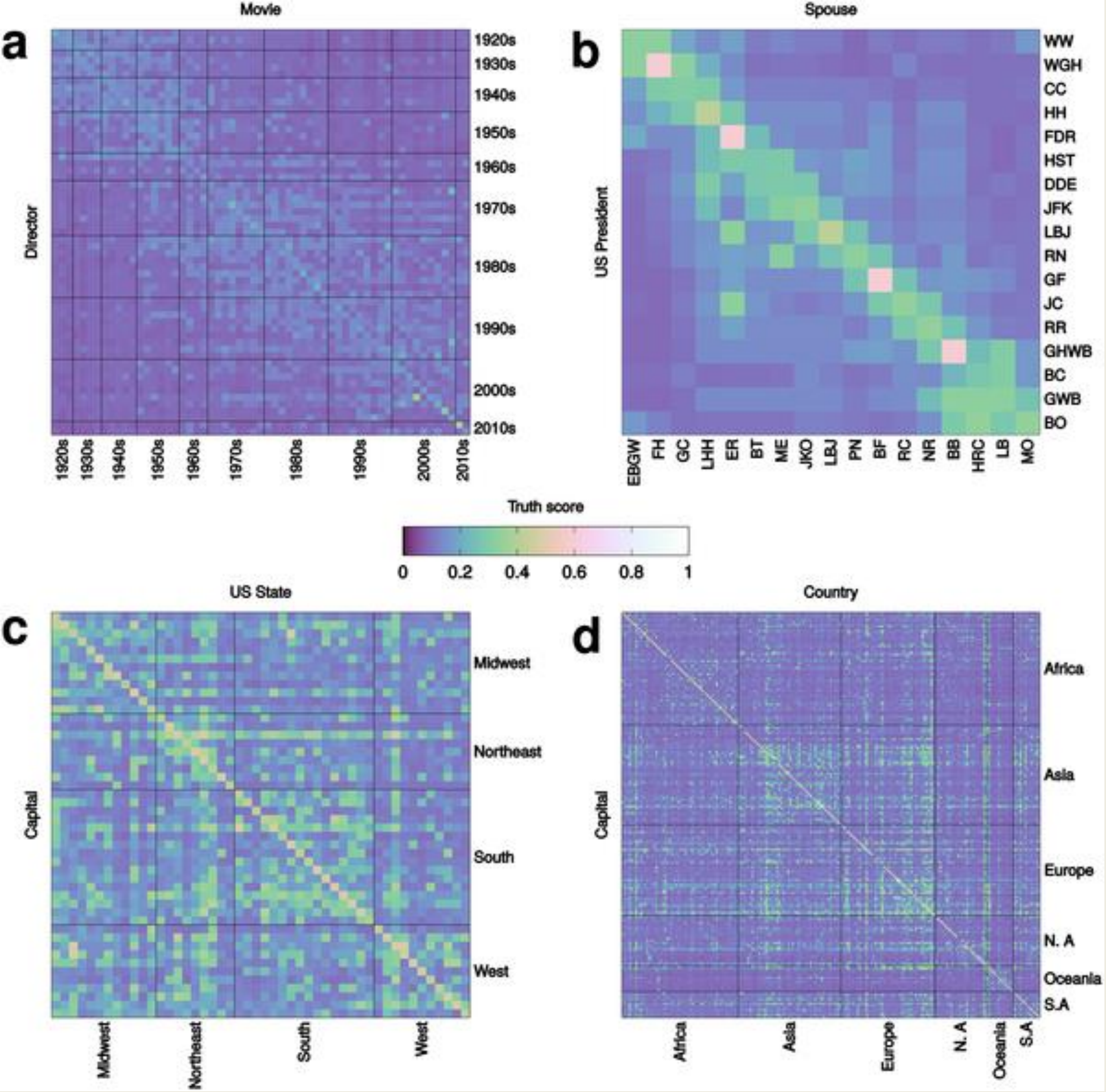
http://cns.iu.edu/docs/netscitalks/2014.10.13-Talk-GiovanniCiampaglia.pdf

Distance function





$$\mathcal{W}_u(P_{s,o}) = \mathcal{W}_u\left(v_1 \ldots v_n\right) = \begin{cases} 1 & \text{if } n=2 \\ \phi\Big(\min_{i=2}^{n-1}\Big\{\log\left(k\left(v_i\right)\right)\Big)\Big\} & \text{if } n>2 \end{cases}$$

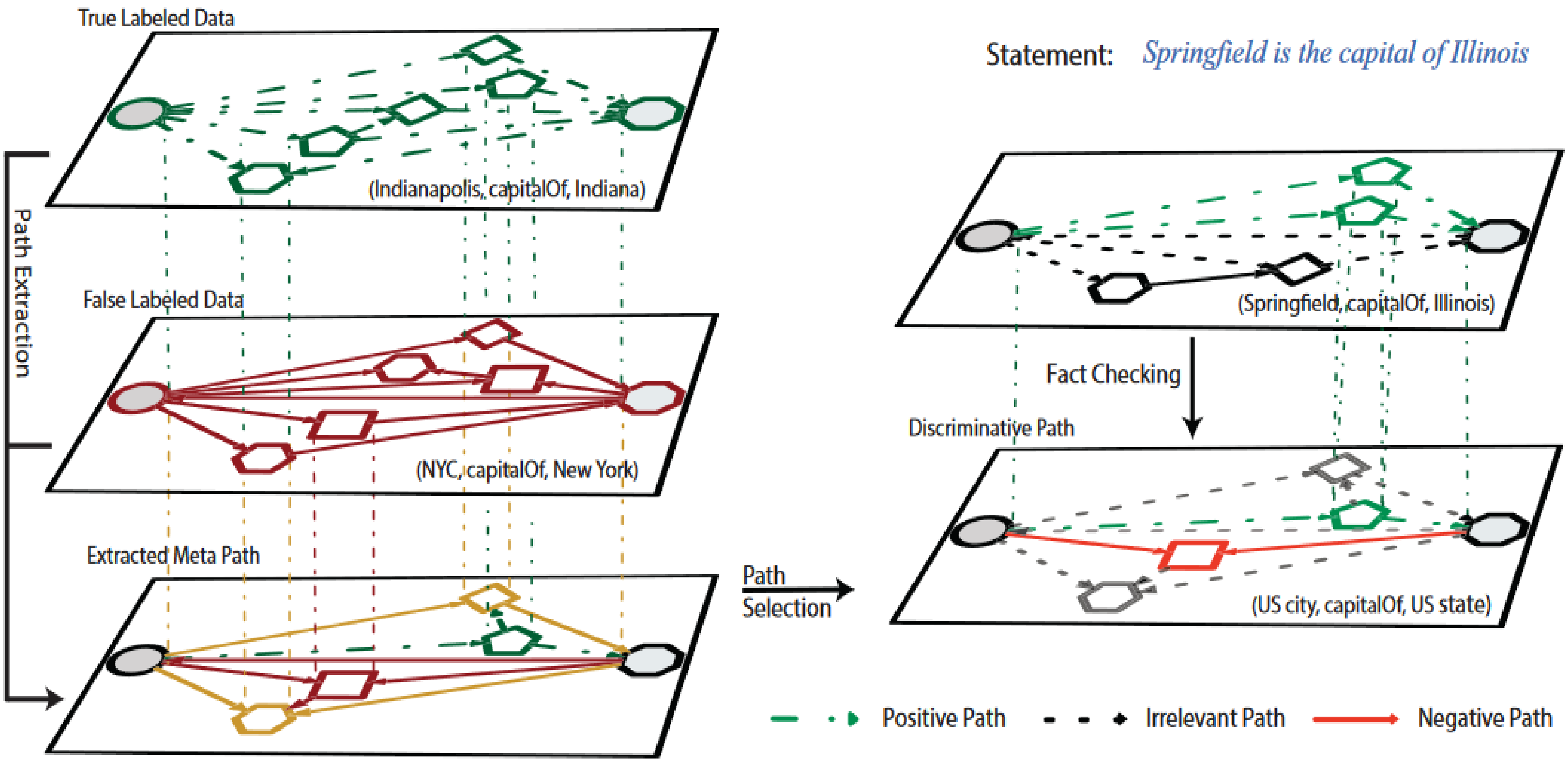


Ciampaglia GL, Shiralkar P, Rocha LM, Bollen J, Menczer F, et al. (2015) Computational Fact Checking from Knowledge Networks. PLOS ONE 10(6): e0128193. https://doi.org/10.1371/journal.pone.0128193

http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0128193

Knowledge Network Approach

- Views this problem as a link-prediction task in a knowledge graph. **[Shi et al., 2015]**



Finding Stream in Knowledge Graph

- Fact checking as a Minimum Cost Maximum Flow Problem [Shiralkar et al., 2017]

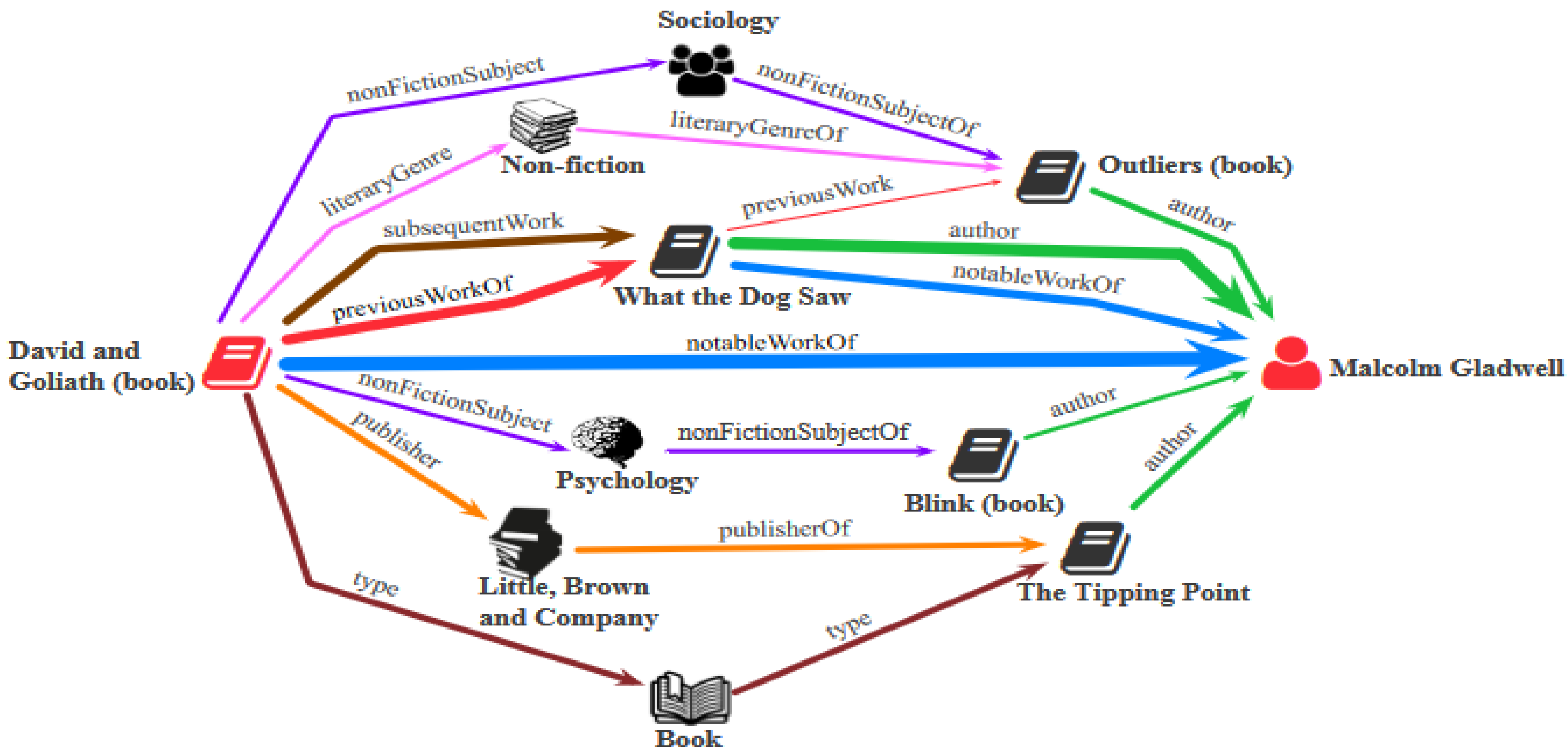
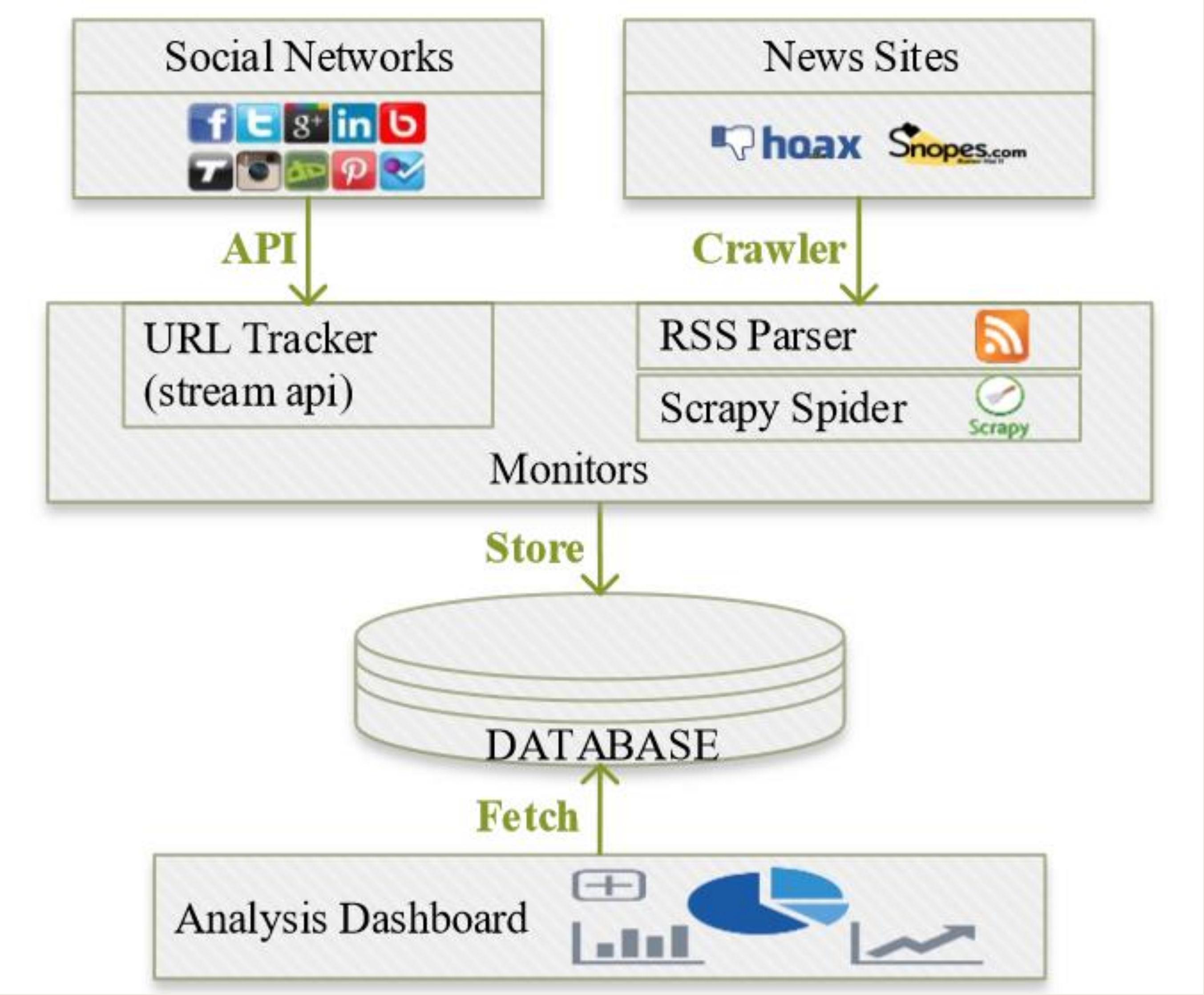


Fig. 1. The best paths identified by Knowledge Stream for the triple (David and Goliath (book), author, Malcolm Gladwell). The width of an edge is roughly proportional to the flow of knowledge through it.

Monitoring Facts and Tracking Misinformation

Hoaxy System Architecture [Shao et al., 2016]



HOAXY Help Stats Data Source OSoMe

Select claims and fact checking articles to visualize.

Select Top 20

3 Visualize

Tweet Share

Soros-Funded Groups Claim Repealing Net Neutrality Is "Racist"
From infowars.com on Nov 27, 2017 — **713** Tweets **3311** Facebook Shares

Flashback: Socialists Developed Net Neutrality to Censor Internet
From infowars.com on Nov 21, 2017 — **420** Tweets **1343** Facebook Shares

YouTube Censors Ajit Pai Net Neutrality Video
From infowars.com on Dec 20, 2017 — **203** Tweets **9** Facebook Shares

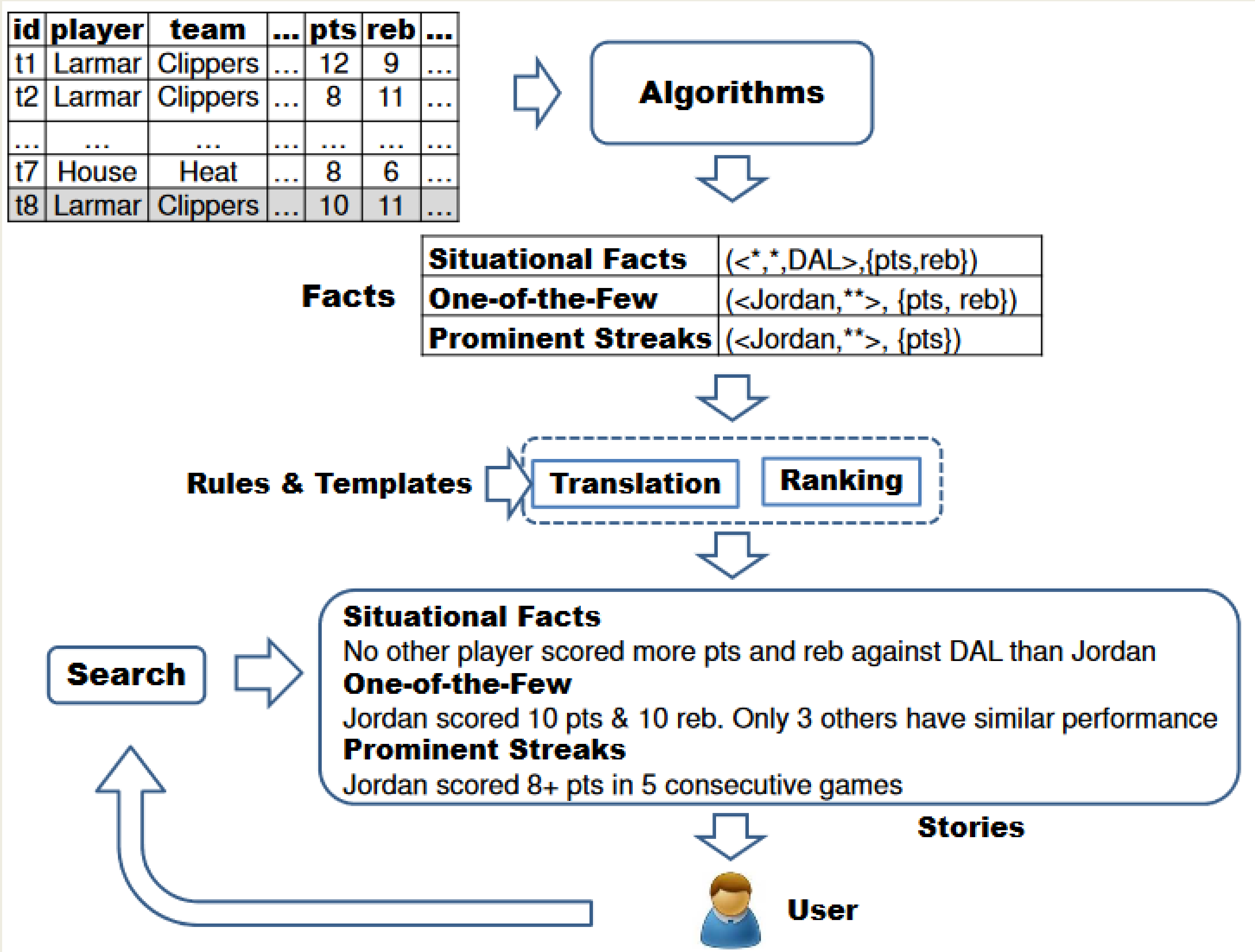
Koch nonprofit president's anti-net neutrality campaign
From opensecrets.org on Dec 13, 2017 — **133** Tweets **1802** Facebook Shares

FACT CHECK: Is Portugal an Example of What Happens Without Net Neutrality?
From snopes.com on Dec 3, 2017 — **126** Tweets **574** Facebook Shares

FCC Votes to Begin Net Neutrality Repeal
From breitbart.com on May 18, 2017 — **118** Tweets **294** Facebook Shares

https://hoaxy.iuni.iu.edu

FactWatcher [Hassan et al., 2014]



FactWatcher NBA 61 Weather Nov 1, 1991 Apr 20, 2005 Speed 1x

Feb 20, 1998

»LIVE UPDATE

[February 20, 1998] Antonio Davis had 12 rebounds, and 2 blocks in the Indiana Pacers' win over the Orlando Magic. This had been done only 6 times before by Antonio Davis. The same (or better) can be said for 18 other performances.

SHOW GRAPHS

SEARCH

FACT TYPE

RANKING

PLAYERS

TEAMS

SEASONS

MONTHS

OPPOSITION TEAMS

62 [March 4, 1992] Dennis Rodman had 34 rebounds in the Detroit Pistons' victory against the Indiana Pacers. No one before had a better or equal performance in NBA history. MORE LIKE THIS

19 [January 13, 1993] In the Detroit Pistons' victory against the Dallas Mavericks, Dennis Rodman became the only player in NBA history to have at least 16 rebounds for 24 consecutive games. MORE LIKE THIS

15 [March 3, 1993] Rony Seikaly had 34 rebounds in the Miami Heat's match against the Washington Bullets. No one before had a better performance in NBA history. MORE LIKE THIS

8 [January 22, 1994] Dennis Rodman had 32 rebounds in the San Antonio Spurs' win over the Dallas Mavericks. This had been done only once before by Dennis Rodman. The same (or better) can be said for 3 other performances. MORE LIKE THIS

http://idir.uta.edu/factwatcher/

TwitterTrail [Metaxas et al., 2015]

Web-based interactive tool that uses an algorithm to measure the trustworthiness of a tweet.

- Who broke the story and when?
- How is the story spreading?
- Who are the main actors of the event?
- How are the communities self-describing?

http://www.twittertrails.com/

Crowdsourced Fact-checking

https://reporterslab.org/fact-checking/

Why fact-checking sites close?

- "Fact-checking is time consuming and therefore expensive."- Nicolas Patte
- Tied to offer impartial service; eschew partisan funding sources.
- Less scalable and more labor intensive.
- Funding is often tied to election cycle.

https://www.poynter.org/news/why-do-fact-checking-sites-close-and-how-can-new-ones-avoid-fate

What's a sustainable model for fact-checking?

- I don't know :)
- But, we **question** ourselves, is the following combination sustainable?

Crowd

\+

Professional

\+

Automation

What role does Crowd play in fact-checking?

- Average fact-checking time is about 7 days.

Action	# Comments	Percentage
Provide Argument	1942	69.56
Post Irrelevant	402	14.40
Seek Clarification	355	12.71
Assign Flair	141	5.05
Check Verifiability	48	1.72

Flair	Avg. # Factual Evidence	Avg. # Opinion/Judgment
Mostly True	3.22	1.83
Half True	3.17	1.67
Confirmed	3.10	1.56
False	2.89	2.59
Mostly False	2.75	1.86
Partisan Bias	2.73	1.90
Please Verify	2.59	2.40
Unverifiable	2.38	3.33

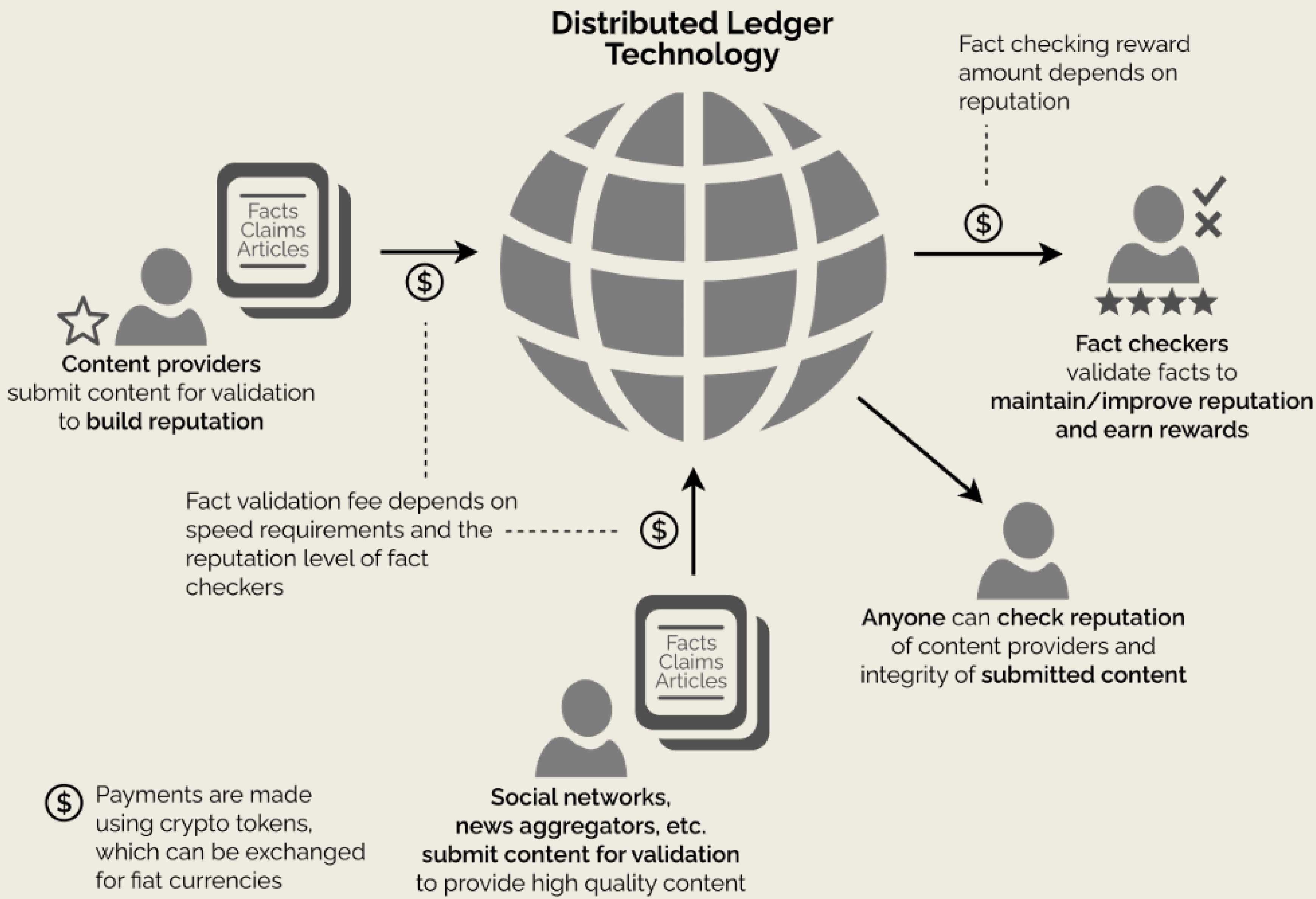
Domain	# Citations
wikipedia.org	298
politifact.com	150
reddit.com	111
washingtonpost.com	91
nytimes.com	78
youtube.com	68
whitehouse.gov	51
factcheck.org	40

Hassan et al., 2017c

Q4: How Automation can help scaling-up?

- Ranking fact-check request by their check-worthiness.
  - ClaimBuster
- Argument Classification
  - Classes: Based on evidence (E), Evidence lacking opinion/judgement (O)
  - Average Accuracy: **82.2%** (+/- 0.17%)
- Stance Detection
- Comments Ranking
- Soliciting arguments from users based on expertise

4Fact.org


Distributed Ledger Technology
Fact checking reward amount depends on reputation
Facts Claims Articles
Content providers submit content for validation to build reputation
Fact checkers validate facts to maintain/improve reputation and earn rewards
Fact validation fee depends on speed requirements and the reputation level of fact checkers
Anyone can check reputation of content providers and integrity of submitted content
Facts Claims Articles
Payments are made using crypto tokens, which can be exchanged for fiat currencies
Social networks, news aggregators, etc. submit content for validation to provide high quality content

Part 4. Postlude

Organizations and Industries

- PolitiFact [www.politifact.com]
- FullFact [www.fullfact.org]
- Factcheck [www.factcheck.org]
- Snopes.com
- Reddit #politicalfactchecking

- Factmata [https://factmata.com]
- Newscheck [http://www.newscheck.com/]
- 4facts [http://4facts.org/]

Connect with other researchers.

- Slack group of Misinfocon [https://misinfocon.slack.com/]
- Google Groups "Combating Fake News: The Science of Misinformation“ https://groups.google.com/group/fakenewssci
- The Credibility Coalition https://meedan.com/credibility-coalition/

Future Direction

- Unified Knowledge Base + Credibility
  - Knowledge base from fact-checks
- Natural Language to Query Formulation
- Question Answering System
  - Aggregation
- Query Similarity

Reference

- **[Shu et al., 2017]** Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22-36.
- **[Hassan et al., 2017a]** Hassan, N., Zhang, G., Arslan, F., Caraballo, J., Jimenez, D., Gawsane, S., ... & Sable, V. (2017). ClaimBuster: the first-ever end-to-end fact-checking system. *Proceedings of the VLDB Endowment*, *10*(12), 1945-1948.
- **[Hassan et al., 2017b]** Hassan, N., Arslan, F., Li, C., & Tremayne, M. (2017, August). Toward automated fact-checking: Detecting check-worthy factual claims by ClaimBuster. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1803-1812). ACM.
- **[Ciampaglia et al., 2015]** Ciampaglia, G. L., Shiralkar, P., Rocha, L. M., Bollen, J., Menczer, F., & Flammini, A. (2015). Computational fact checking from knowledge networks. PloS one, 10(6), e0128193.
- **[Shi et al., 2015]** Shi, B., & Weninger, T. (2015). Fact Checking in Large Knowledge Graphs-A Discriminative Predicate Path Mining Approach. arXiv preprint arXiv:1510.05911.
- **[Shao et al., 2016]** Chengcheng Shao, Giovanni Luca Ciampaglia, Alessandro Flammini, and Filippo Menczer (2016). Hoaxy: A Platform for Tracking Online Misinformation. In Proceedings of the 25th International Conference Companion on World Wide Web (WWW '16 Companion). Pages 745-750.

Reference

- **[Hassan et al., 2014]** Hassan, N., Sultana, A., Wu, Y., Zhang, G., Li, C., Yang, J., & Yu, C. (2014). Data in, fact out: automated monitoring of facts by FactWatcher. *Proceedings of the VLDB Endowment*, *7*(13), 1557-1560.
- **[Metaxas et al., 2015]** Metaxas, P. T., Finn, S., & Mustafaraj, E. (2015, February). Using twittertrails. com to investigate rumor propagation. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing* (pp. 69-72). ACM.
- **[Hassan et al., 2017c]** Naeemul Hassan, Mohammad Yousuf, Md Mahfuzul Haque, Javier A. Suarez Rivas, Md Khadimul Islam. Towards A Sustainable Model for Fact-checking Platforms: Examining the Roles of Automation, Crowds and Professionals. In Proceedings of the 2017 Computation+Journalism Symposium (C+J2017), Illinois, October, 2017.
- **[Shiralkar et al., 2017]** Shiralkar, P., Flammini, A., Menczer, F., & Ciampaglia, G. L. (2017). Finding streams in knowledge graphs to support fact checking. *arXiv preprint arXiv:1708.07239*.