Combating Fake News: A Data Management and Mining Perspective

Laks V.S. Lakshmanan←, Michael Simpson←, and Saravanan Thirumuruganathan→

Disclaimers

- What this tutorial is not about:
- theories of fake news
- economic impact
- psychological aspects, social (media) context
- "completeness".
- …

- What we plan to cover

- Possible sensibilities

Outline

- Primer on "Fake News"
- Some Computational Problems
- Propagation
- Detection
  - ML based approaches
  - DB approaches
- Mitigation & Intervention
- Future Challenges & Opportunities

“Fake News” Primer

Some terminology

- “Fake News” comes in many forms

- “Fake” articles / images / videos
- Misinformation and Disinformation
- False / Misleading Claims
- …

FN Definition

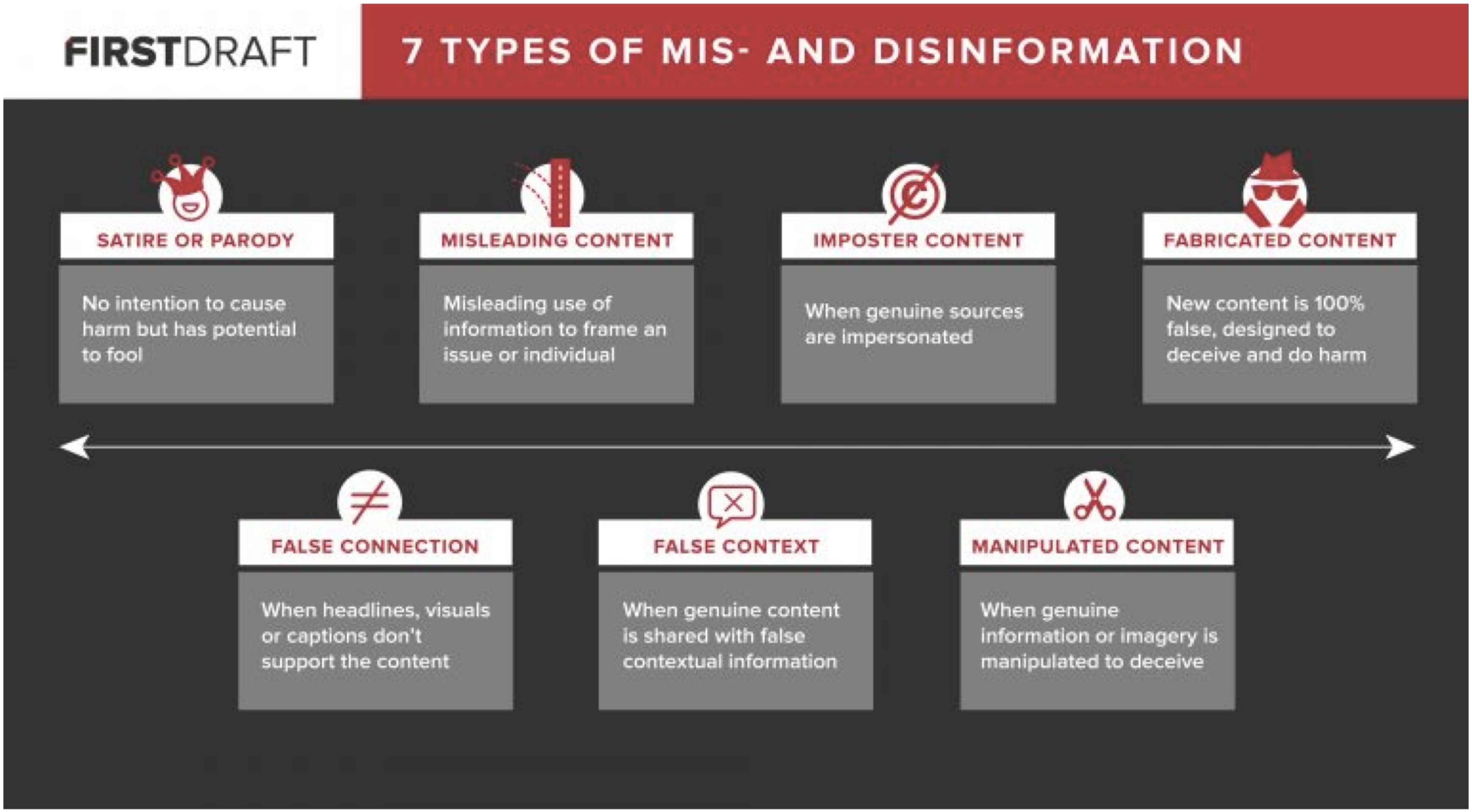
- Abused and misused term

- Different kinds of untruth or misleading info.[§] that is often referred to as "fake news"

- Many prior definitions: e.g., "Internationally and verifiably false" (Allcott et al. 2017).
- **Tutorial focus:** detecting a subset of specific forms of such bad content, modeling their diffusion, detection, and their mitigation & intervention.

[§]As well as some completely genuine news! ☹

A Taxonomy



Fake News. It's complicated.

An Alternative Taxonomy

- **Satire**: no malicious intent; entertainment value (e.g., The Onion, Andy Borowitz: The New Yorker, ...).

[Guo and Vargo, *Communications Research* 2018].

An Alternative Taxonomy

- **Selective disclosure, cherry picking facts** -- some intention to mislead or advance agenda:
  - e.g.1 (*structured data*): Rudy Giuliani's claim "adoptions went up 65 to 70 percent" in NYC "when I was the mayor."
  - true on surface: 1996-2001 vs 1990-1995.
  - Giuliani was mayor 1994-2001.
  - however, from term 1 (1994-97) to term 2 (98-'01) adoptions went down by 1%.

[Wu, Agrawal, Li, Yang, and Yu. *TODS* 2017.]

An Alternative Taxonomy

- e.g.2: zooming in to make a point.

An Alternative Taxonomy

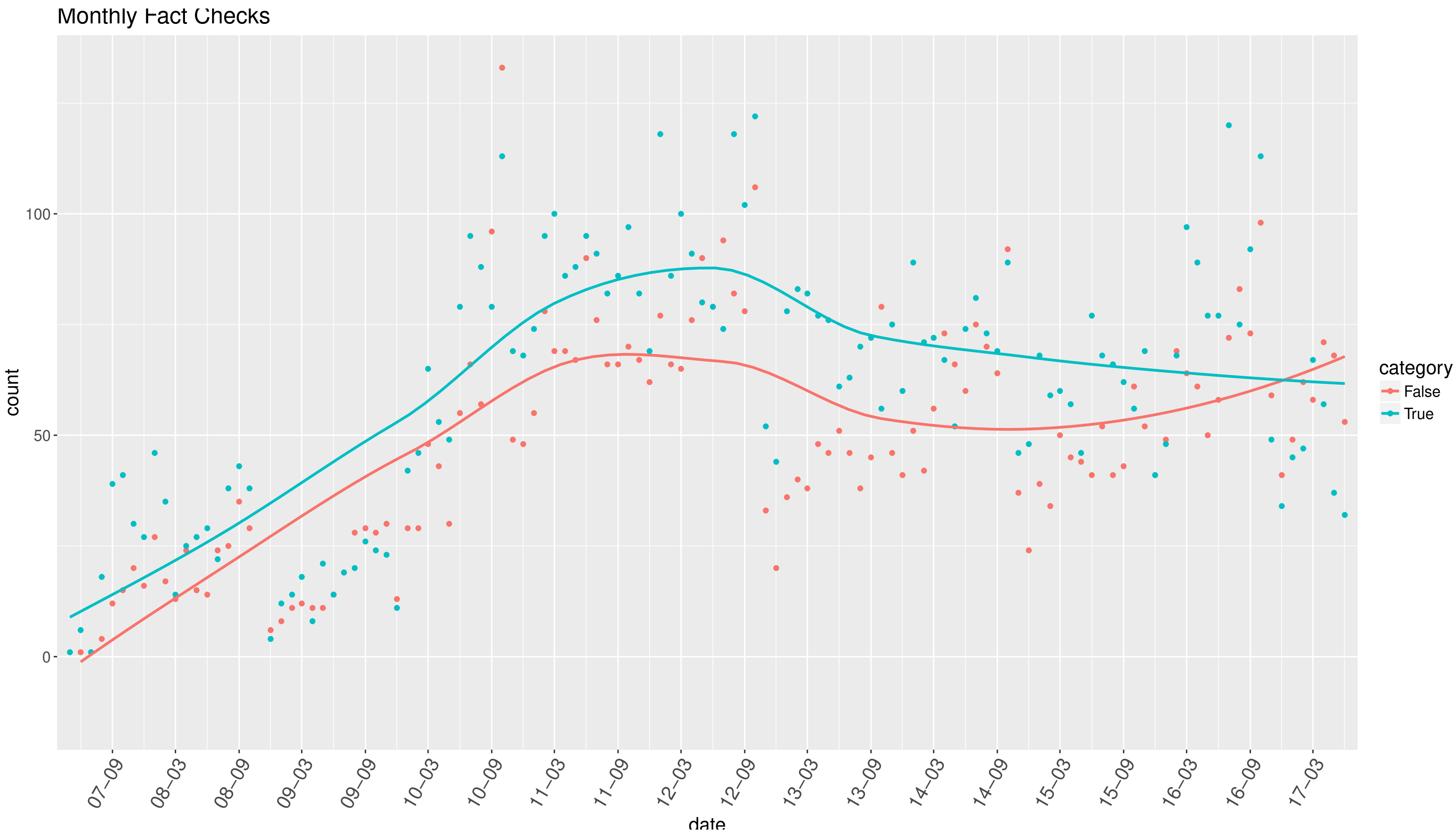
- **make false connections** to advance *conspiracy theories*: e.g., add additional facts/observations (coincidental) to promote CTs.
- **imply false context** to story (image/video) to push a false narrative: e.g., snowfall somewhere as "evidence" against global warming.
- **manipulate** photo/story/facts to paint false picture: e.g., edited video of Nancy Pelosi; climate analytics with different start dates.

An Alternative Taxonomy

- **Complete Fabrication** (usually easier to detect than subtle distortions): e.g., “As the Telegraph’s Brussels correspondent between 1989 and 1994, he invented a self-serving journalistic genre that set a poisonous tone for British EU reporting” *The Guardian*.

- **Impostor – make-believe sites**: make site look and feel authentic and real.

Growth in Fake News



Monthly Fact-checks by Politifact

Impact of FN

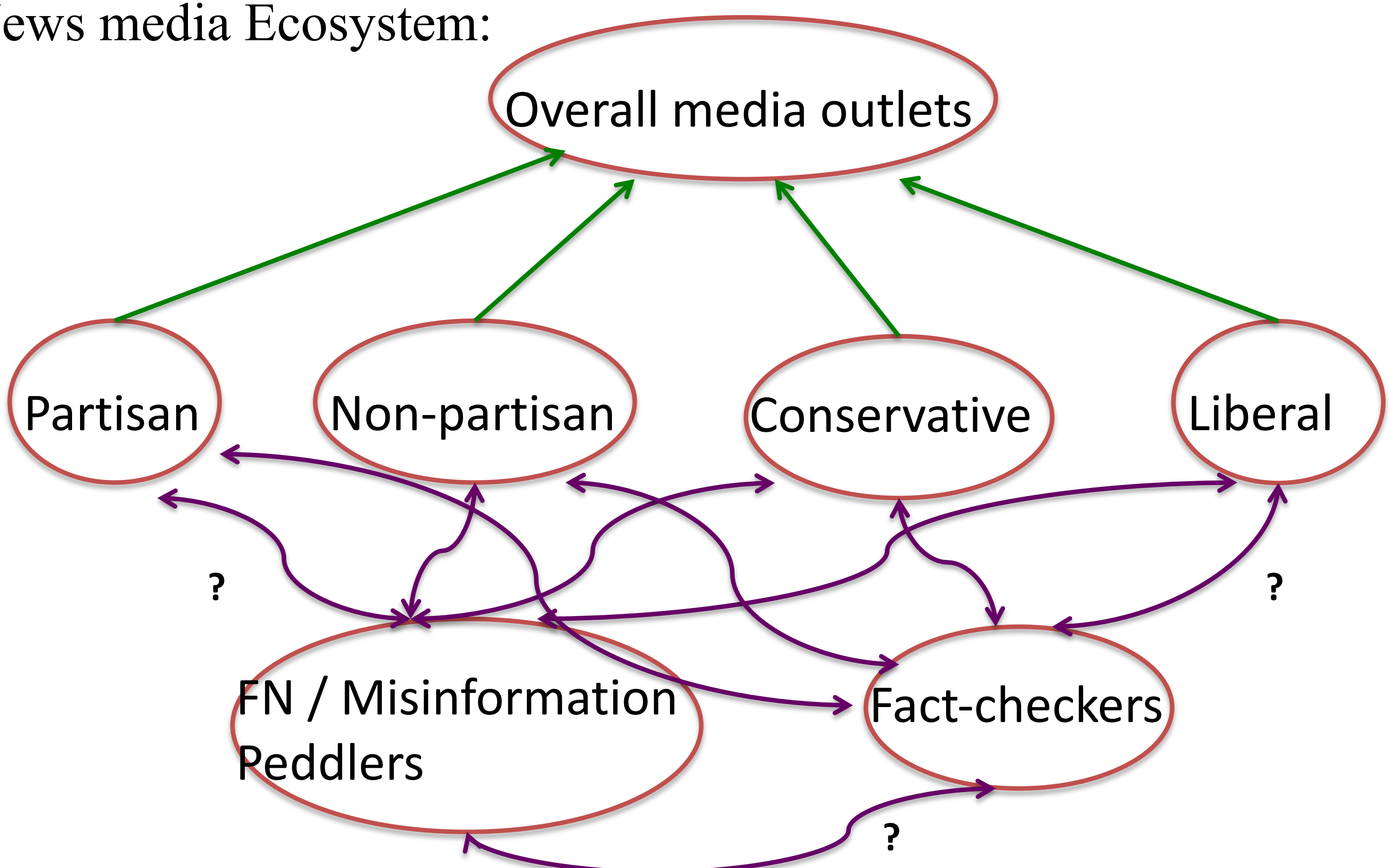
Recent events have amplified the effects of FN:

- Social media, virality, polarization, filter bubbles.
- impact on news media ecosystem, not just on end user (aka consumer).

[Guo and Vargo *Communication Research* 2018].

News Media Ecosystem

News media Ecosystem:



News Media Ecosystem

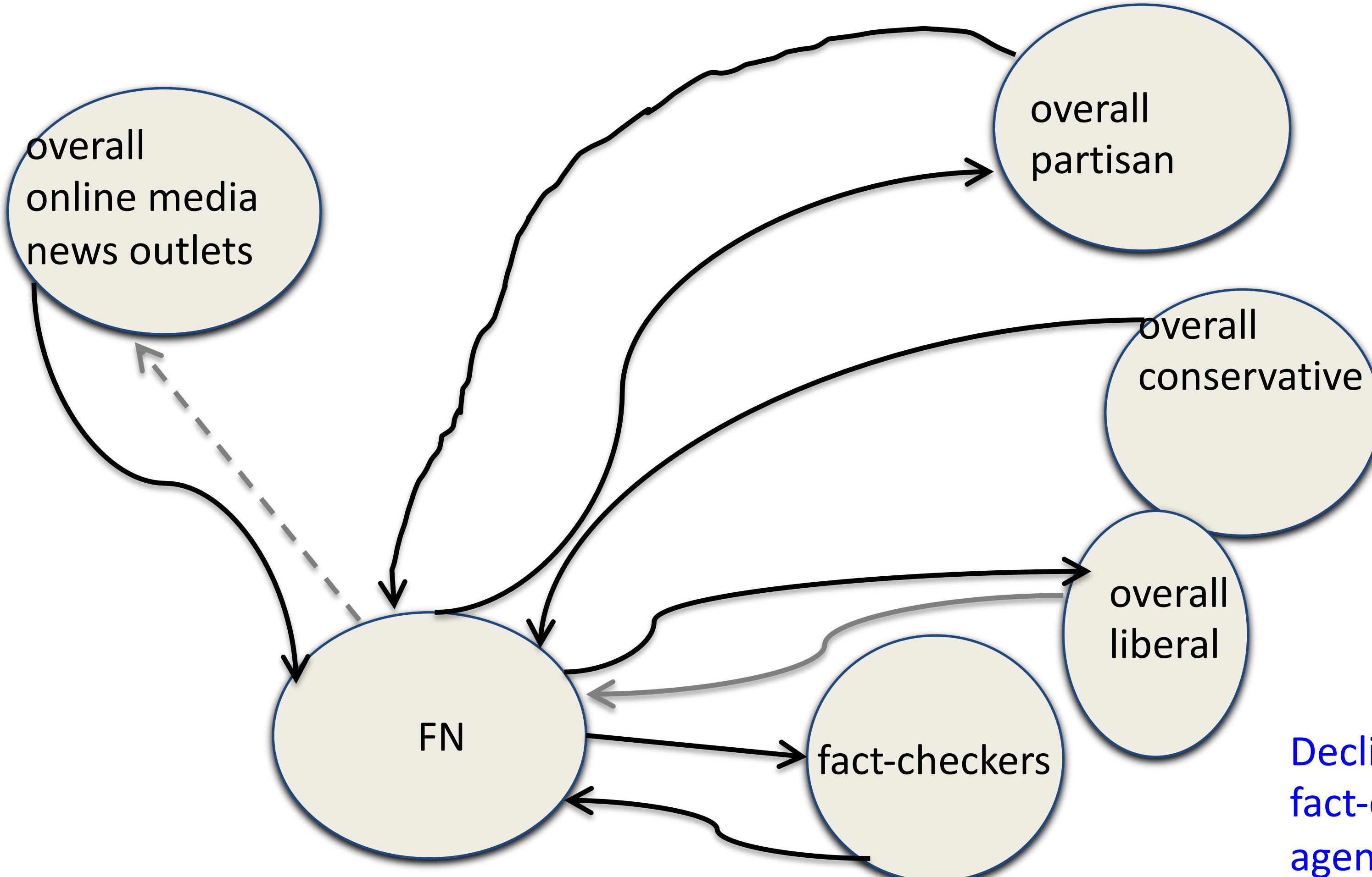
- Elite Mainstream media: normally regarded as opinion leaders.
- smaller outlets -- followers.
- Social platform taking the lead on covering certain stories.

Intermedia Network Agenda Setting.

- Period studied – 2014-2016.

[Guo and Vargo *Communication Research* 2018].

News Media Ecosystem Findings



Declining influence of fact-checkers on agenda of fact-based news orgs!

Based on online media landscape 2014—2016.
Caveat: Our Schematic oversimplified.

2. Some Computational Problems

Is this News Genuine?

- **Problem P1**: Given a repository of real and "fake" news articles, and an article *A*, find if *A* is real or "fake".
  - what metadata is available?
  - propagation patterns?
  - unsupervised, semi-supervised, supervised?

perhaps we can simultaneously grade sources and articles (and perhaps commentators) leveraging all available signals.

Fact Checking Claims – Simple

- **Problem P2**: Given a claim C and a collection $\mathcal{A}$ of articles, determine if C is true or false.
  - C is a simple factual assertion.
  - collection $\mathcal{A}$ is assumed to contain relevant articles.
  - different shades of truth in place of just true/false.
  - **subproblem**: determine if an article A supports or refutes a claim C, is related or unrelated to it.
    - related to *stance detection.*

Fact Checking Claims – Quantitative

- **Problem P3:** Given a claim C and a collection $\mathcal{A}$ of articles, determine if C is true or false.
  - C is an aggregate statement.
  - room for cherry picking, by careful choice of window (could be geo or time) that C applies to.
  - Of course, outright falsehood is (always) possible and is easier to detect than cherry picked assertions.

Querying Knowledge Graphs – Simple

- **Problem P4:** Given a claim C and a knowledge graph $\mathcal{G}$, determine if C is true or false.
  - C is a simple factual assertion.
  - KG $\mathcal{G}$ is assumed to contain relevant facts.
  - different degrees of truth.

Querying Knowledge Graphs – Quantitative

- **Problem P5:** Given a claim C and a collection knowledge graph $\mathcal{G}$, determine if C is true or false.
  - C is an aggregate statement.
  - KG $\mathcal{G}$ is assumed to contain relevant facts.
  - how do you query a KG for aggregate claims?

Mitigation

Problem P6: Given a misinformation campaign, how to effectively counter it?

- propagation model?
- objective of counter campaign?
- before or after misinfo. campaign is underway?

Intervention

Problem P7: Given a misinformation campaign, how to intervene with the content's propagation?

- soft or hard?
- network or content?
- nodes or edges?

3. Propagation of Fake News

Why study Fake News Propagation?

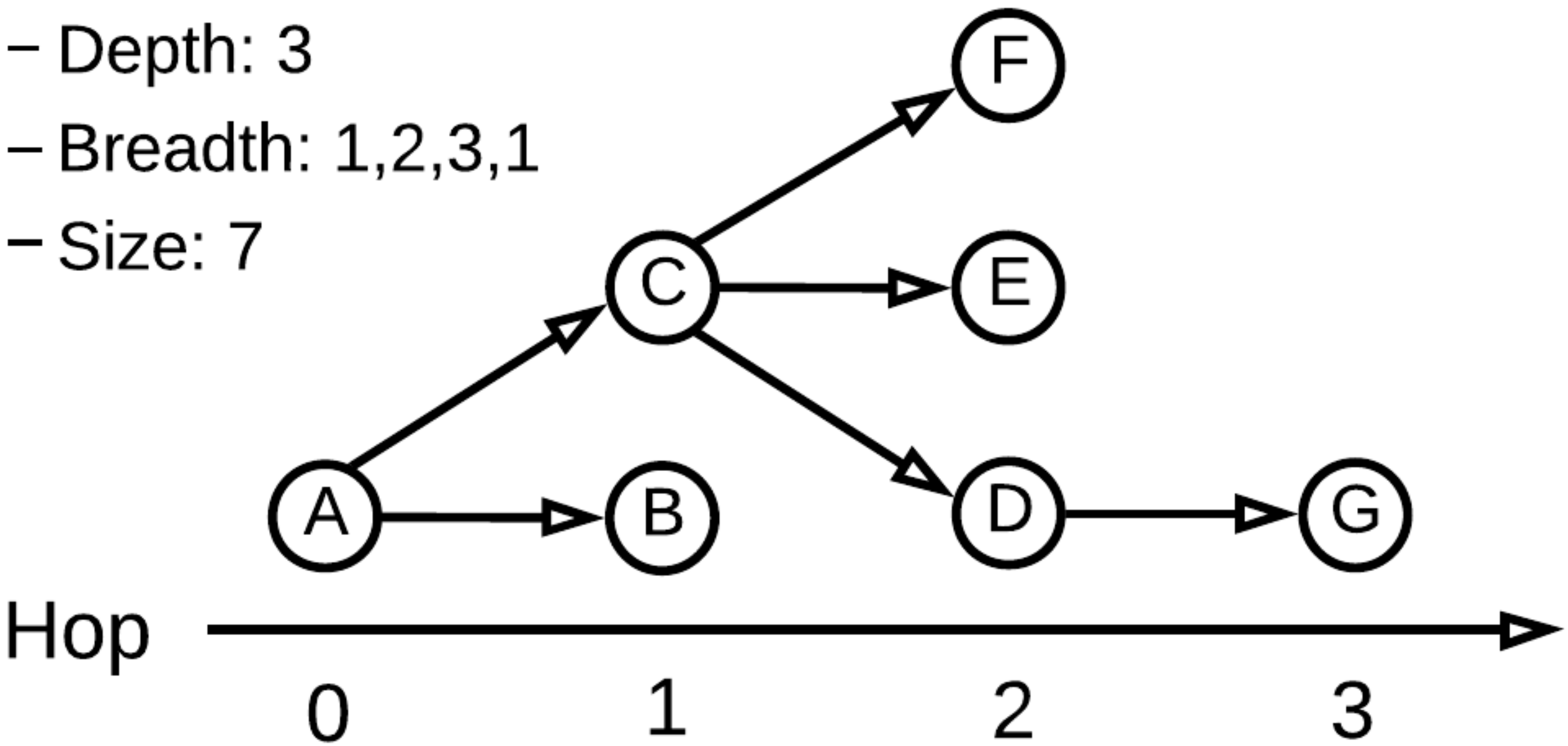
- Understand difference between real and fake news propagation

- Could be used for detection and mitigation

Fake News Cascades

- Most common representation to study propagation
- Tree like structure
  - Root node : initial poster
  - Other nodes: Subsequent posters/retweeters
  - Directed edge between poster and reposter
  - Additional metadata such as timestamp included as necessary

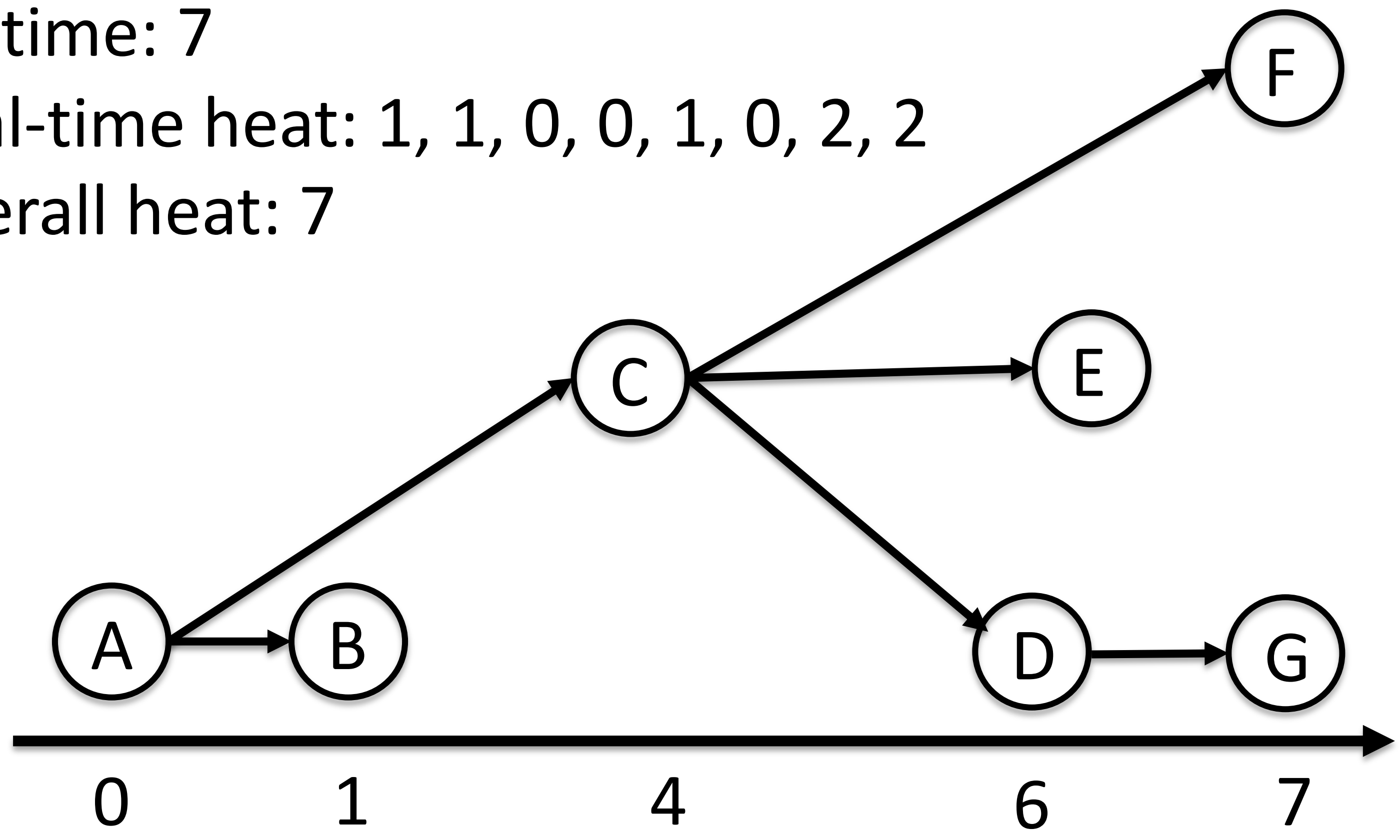
Hop based Fake News Cascades



[Zhou, Zafarani. *arXiv*. 2018.]

Time based Fake News Cascades

- Lifetime: 7
- Real-time heat: 1, 1, 0, 0, 1, 0, 2, 2
- Overall heat: 7



[Zhou, Zafarani. *arXiv*. 2018.]

Empirical Patterns

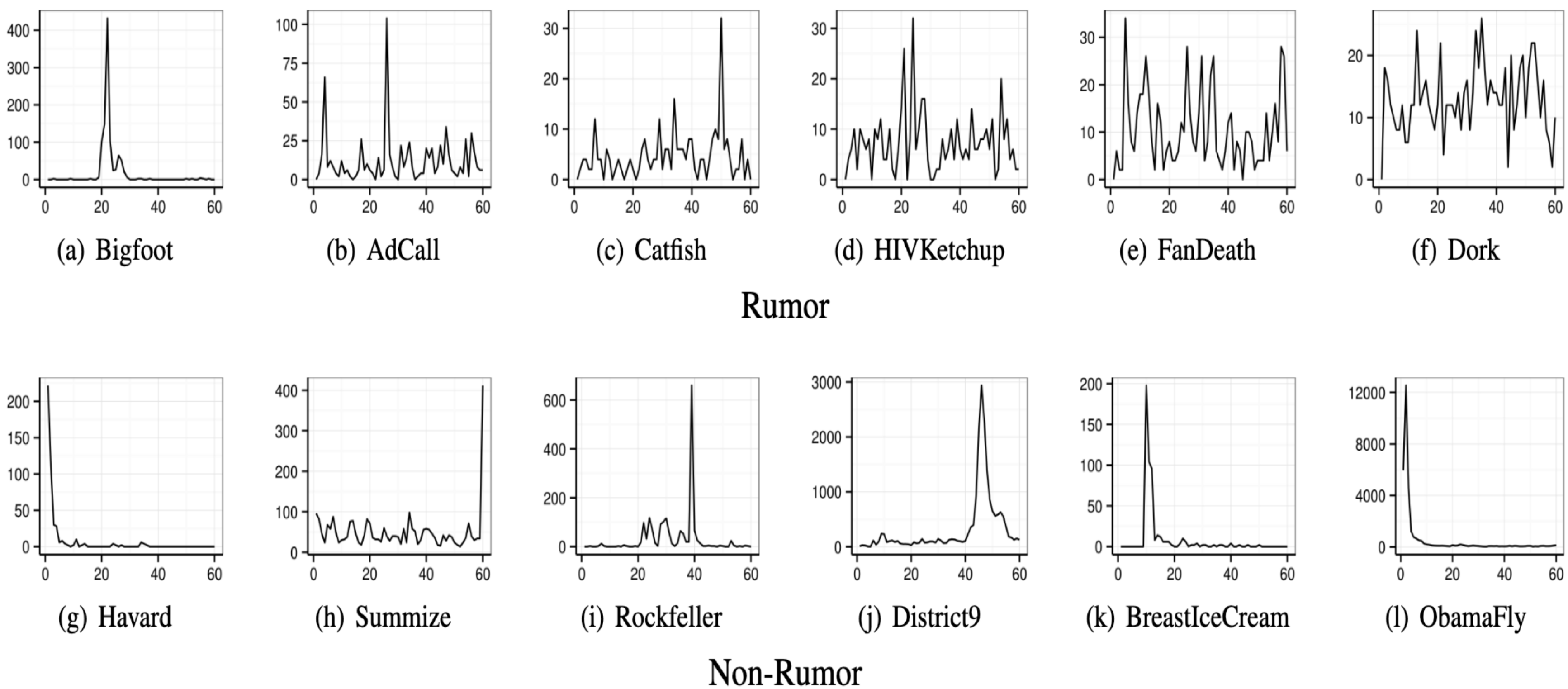
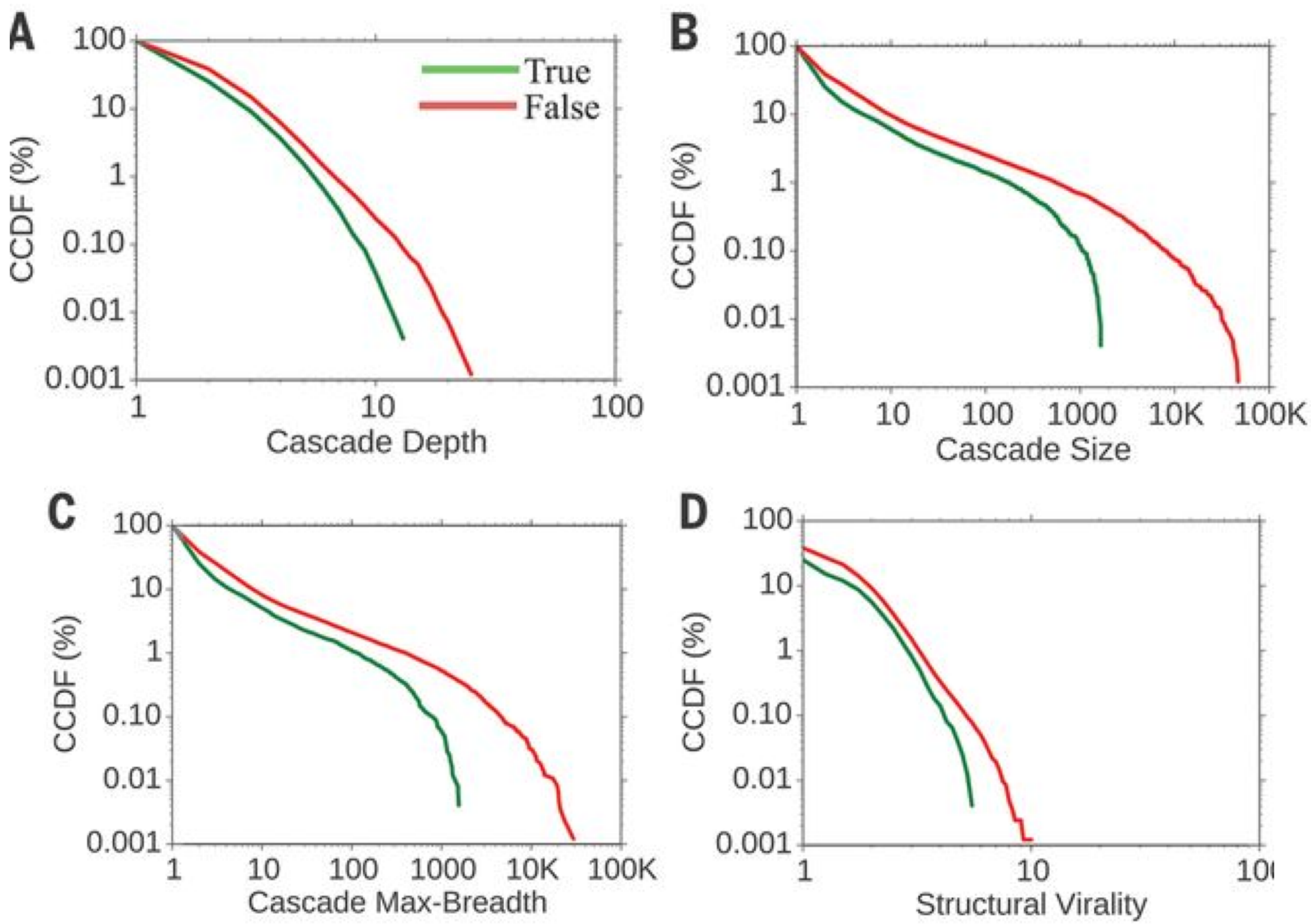


Figure 1: Examples of extracted time series, with x-axis as days and y-axis as the number of tweets on the topic.

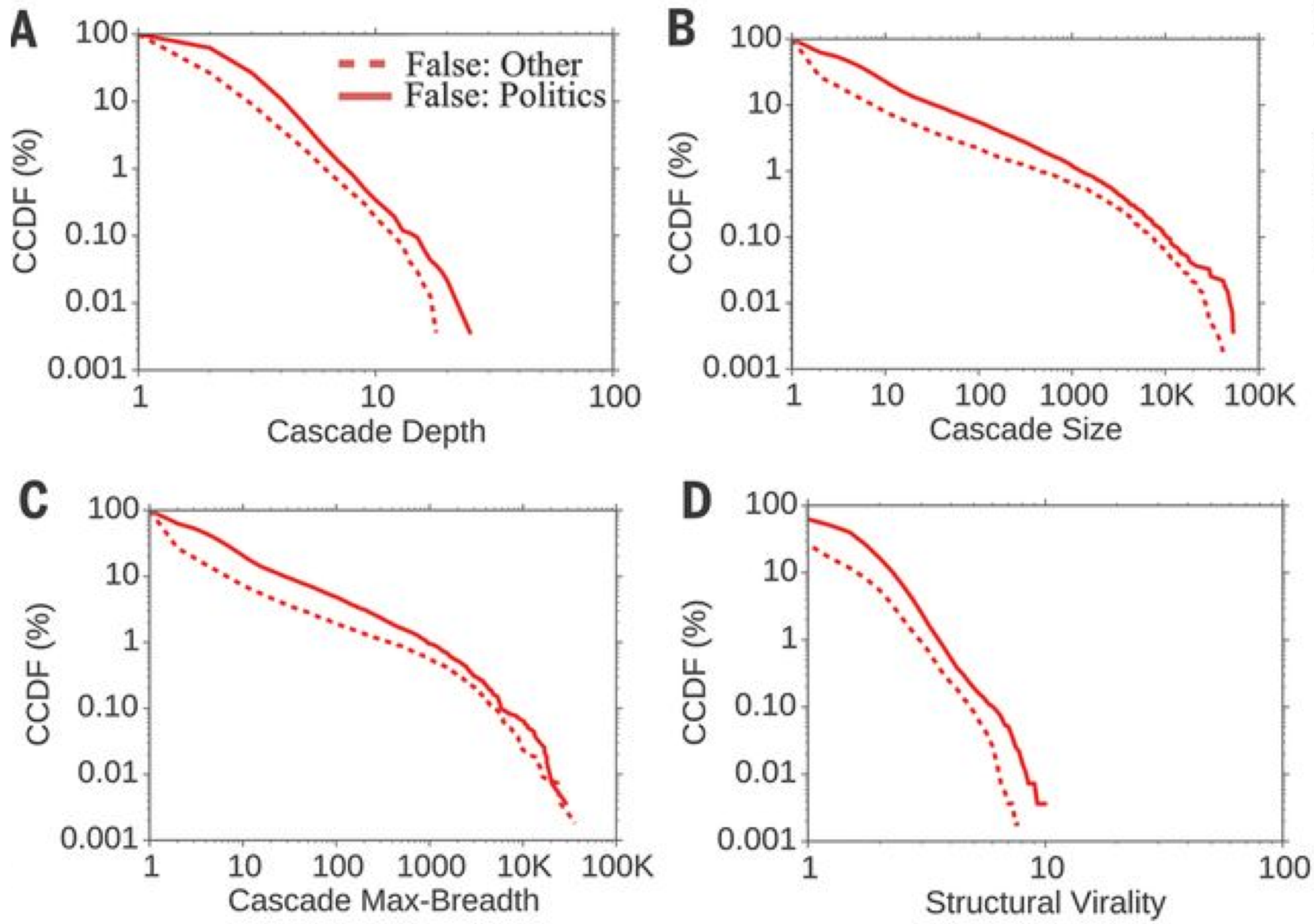
[Kwon et al. *ICDM* 2013.]

Empirical Patterns



[Vosoughi, Roy, Aral. *Science*. 2018.]

Empirical Patterns

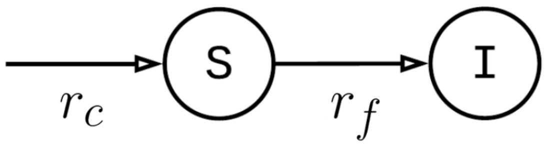


[Vosoughi, Roy, Aral. *Science*. 2018.]

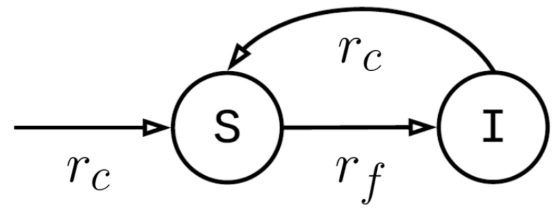
Modeling Fake News Propagation

- So far: Quantitative analysis of propagation
- Need: Mathematical models for quantifying and predicting the propagation
- How can we reuse ”growth” models from other communities?

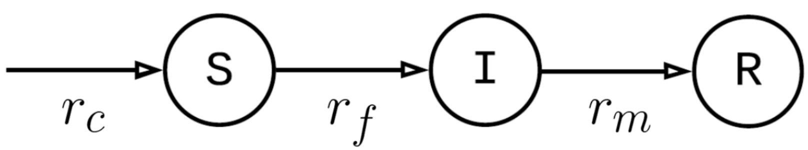
Epidemic Diffusion Model



SI model



SIS model



SIR model

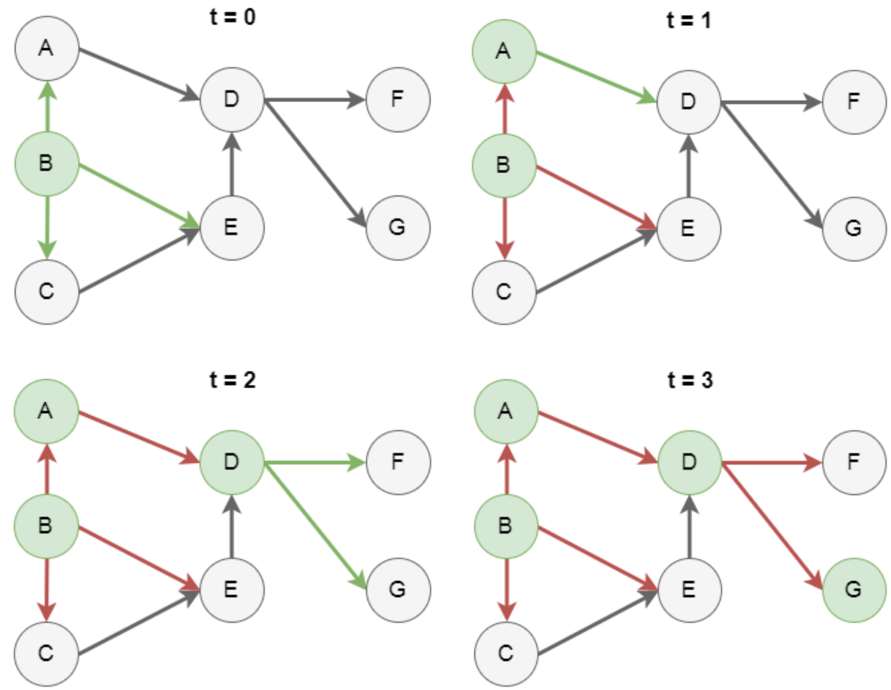
S: Susceptible	r_c: Contact rate
I: Infected	r_f: Infection rate
R: Recovered	r_m: Recovering rate

[Zhou, Zafarani. *arXiv*. 2018.]

IM: Independent Cascade Model

Diffusion of information under IC occurs in a series of rounds:

1. Activate seed set
2. in each round, newly active nodes have single chance to activate inactive neighbours
3. Use influence probabilities on edges to resolve activations
4. Active nodes do not de-activate



seed = *B*

[Kempe et al. *KDD* 2003] [Image: Gursoy, Durahim. *arXiv*. 2018]

4. Detection of Fake News

Detection of Fake News

- ML based approaches
  - Feature engineering (content, credibility, network, propagation)
  - Training a classifier
- DB based approaches
  - Richer set of possibilities
  - Focus: Fake news detection by fact checking

Fact Checking

- Computational Problem P2
- Input: a factual statement
  - whose correctness could be verified
- Output: verdict on correctness of the statement

4a. ML based Detection of Fake News

Supervised ML Approaches

- Related to computational problems P1 and P2

Steps

- Dataset collection and Feature Engineering
- Training a model from labeled data
- Making predictions in the real-world

Issues in Supervised Approaches

- Training data is often small
- Expensive to get accurate labels
- Good feature engineering is often very challenging
- Dataset is often skewed/unbalanced
- Asymmetric cost for misclassification

Feature Engineering

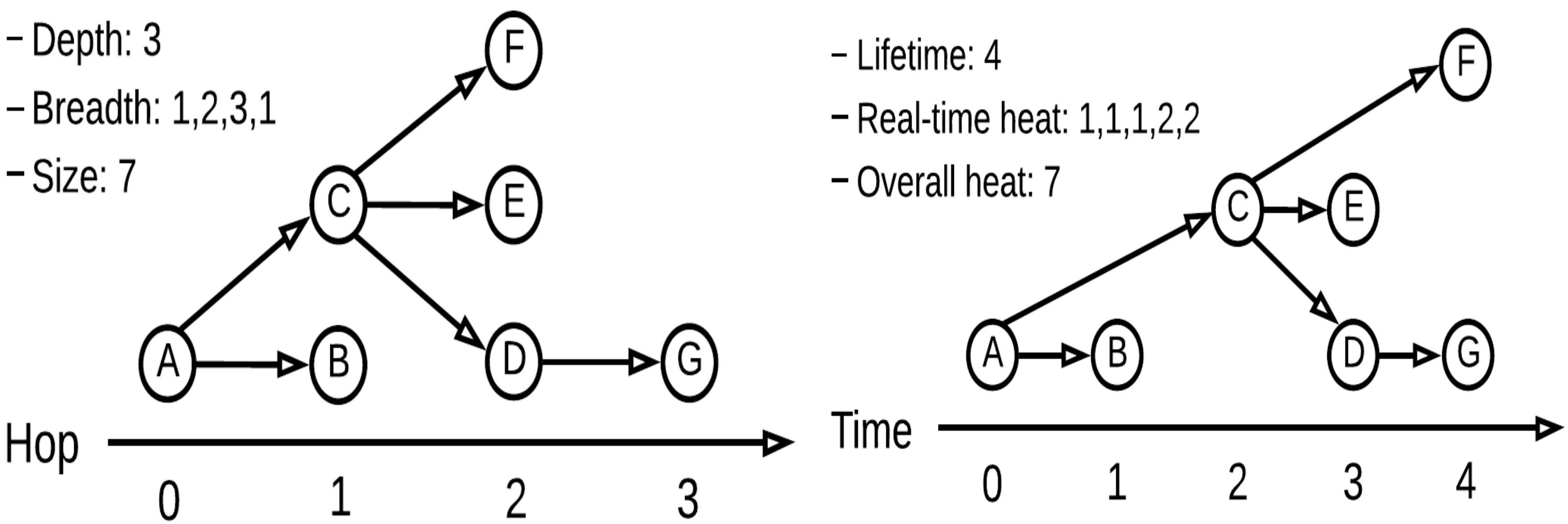
- Poster based
- Network based
- Content based

Content based Features

- Key feature categories
  - Quantity, Complexity , Uncertainty, Sentiment, Typographical, Readability
- Early Approaches
  - Perform content feature engineering
  - Train a classifier and use it for predictions

[Zhou, Zafarani, Shu, et al. *WSDM*. 2018.]

Network-based Features



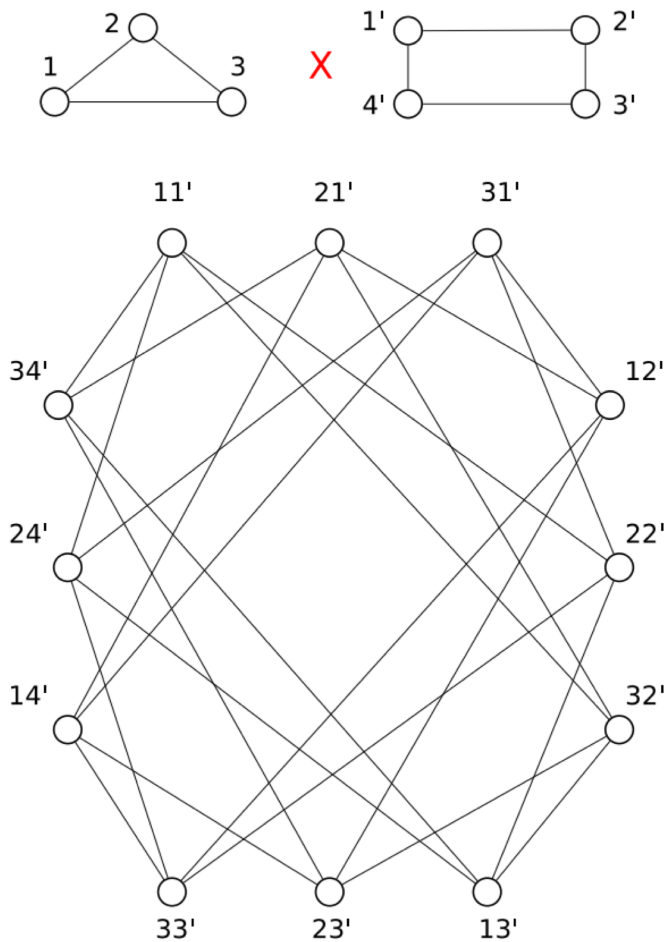
[Zhou, Zafarani. *arXiv*. 2018.]

Propagation-based Approaches

- Cascade features:
  - breath, depth, virality, time
  - Text of original tweet; retweets; replies;
- Cascade similarity via graph kernels
  - If a cascade is very similar to previous fake news cascades, it is probably fake
- Hybrid features
  - Semantic features such as topics and sentiments
  - User roles such as opinion leader or normal user
  - Approval, sentiment, and doubt scores among user posts

[Wu, Yang, Zhu. *ICDE* 2015.]

Cascade Similarity via Graph Kernels



[Vishwanathan, Schraudolph, Kondor et al. *JMLR* 2010.]

Joint Probabilistic Modeling

- Past approaches: build classifier based on single set of features (user, network, content based etc)

- Current approaches:
  - Use all feature classes
  - Understand their interaction and do joint probabilistic modeling

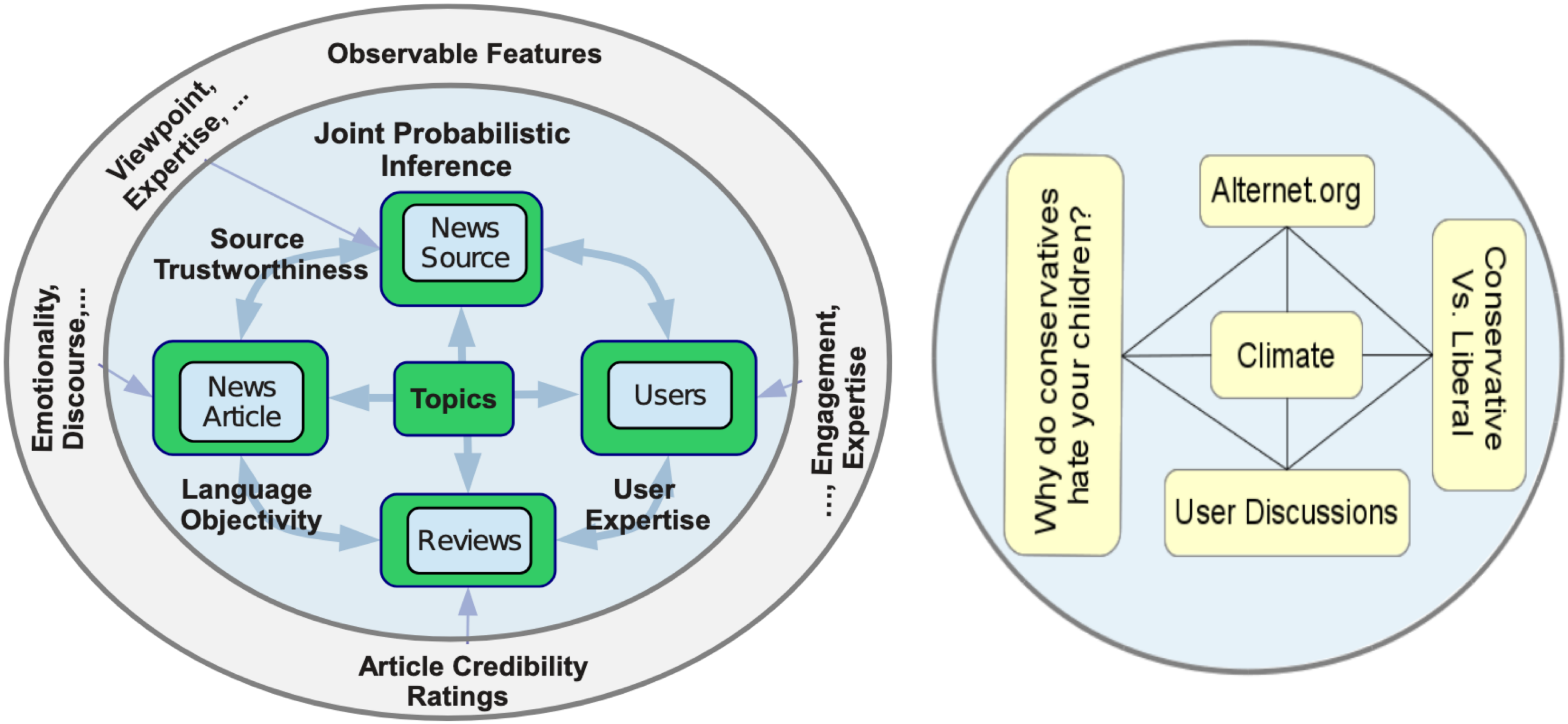
Joint Probabilistic Modeling

Problem Statement: Given a set of news sources generating news articles, and users reviewing those articles on different qualitative aspects with mutual interactions, identify

- Highly credible news articles
- trustworthy news sources
- expert users who perform the role of "citizen journalists" in the community.

- Related to Computation problem P1

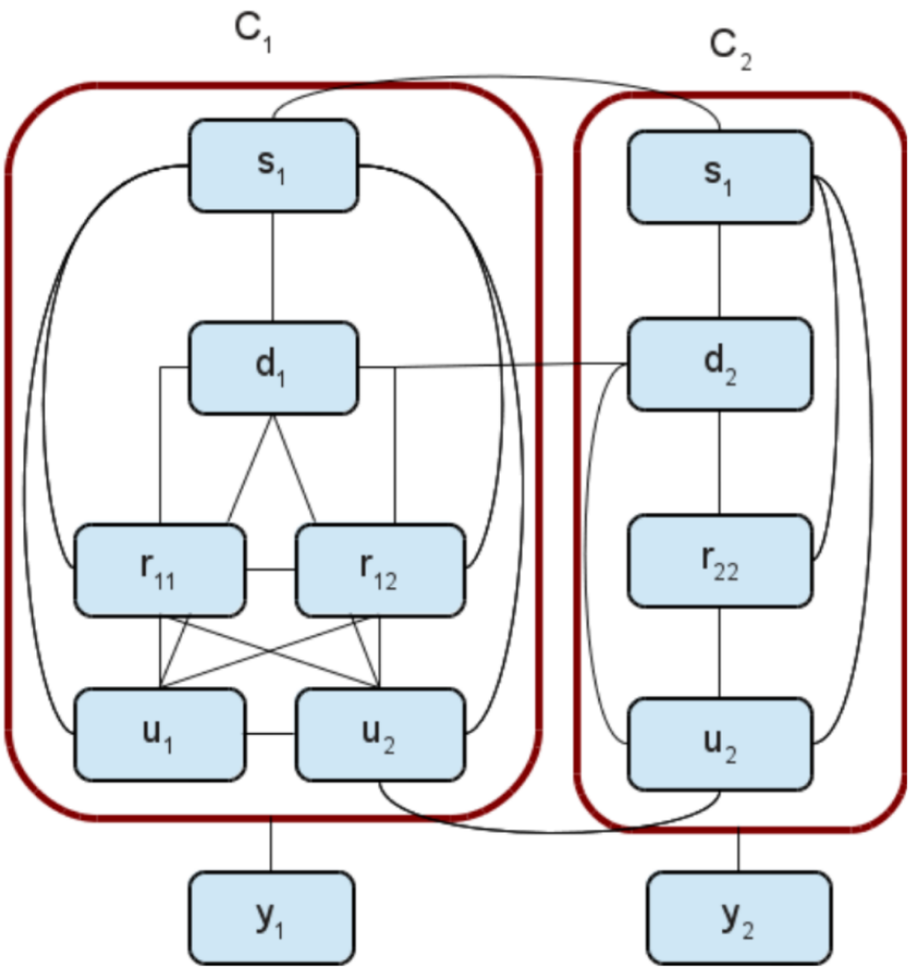
Joint Probabilistic Modeling



- Interactions between source trustworthiness, article credibility, language objectivity, and user expertise.

[Mukherjee, Weikum. *CIKM* 2015.]

Probabilistic Graphical Model



Communities

News Sources

News Articles

Reviews on articles

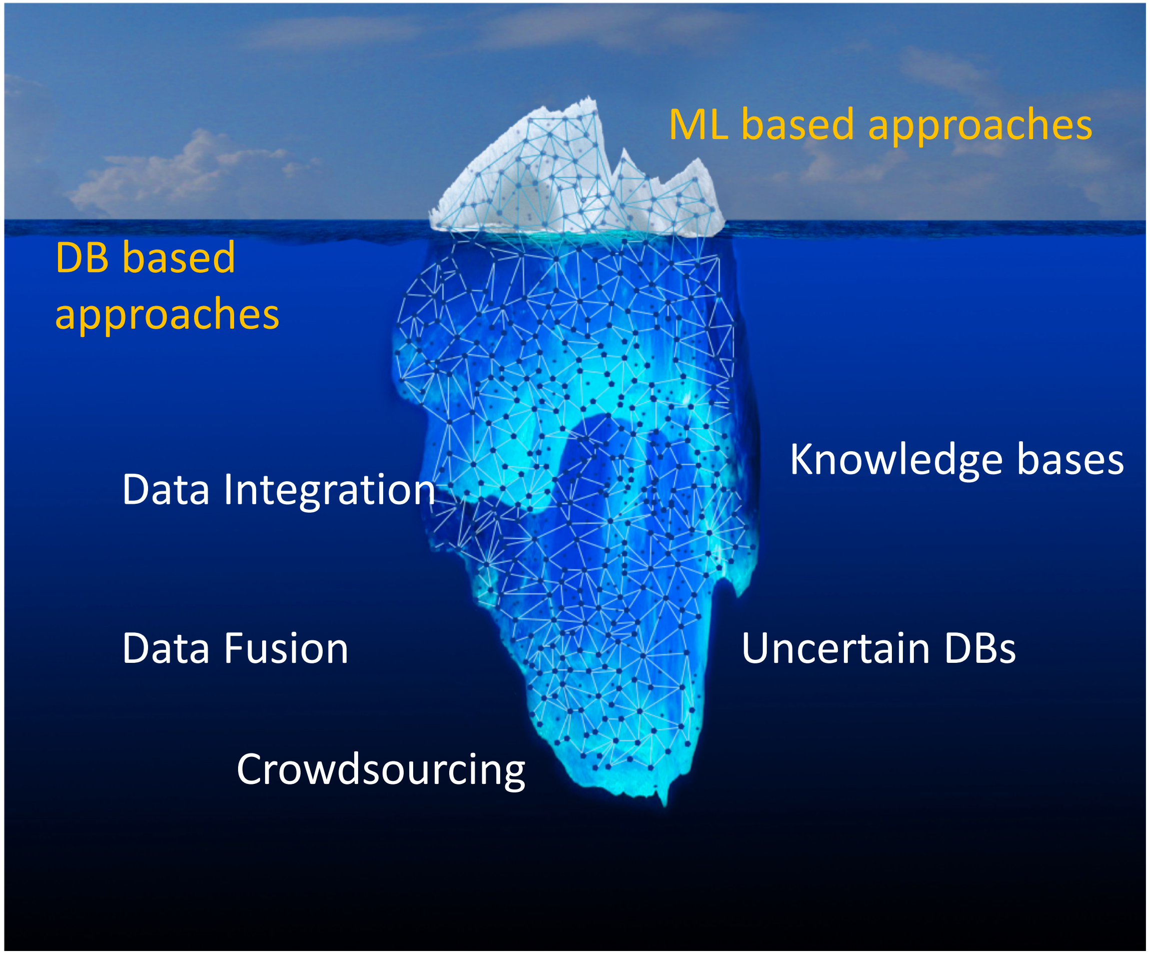
Users

Credibility Ratings

[Mukherjee, Weikum. *CIKM* 2015.]

4b. DB based Detection of Fake News

Detection of Fake News



Fact Checking

Two Step Process:

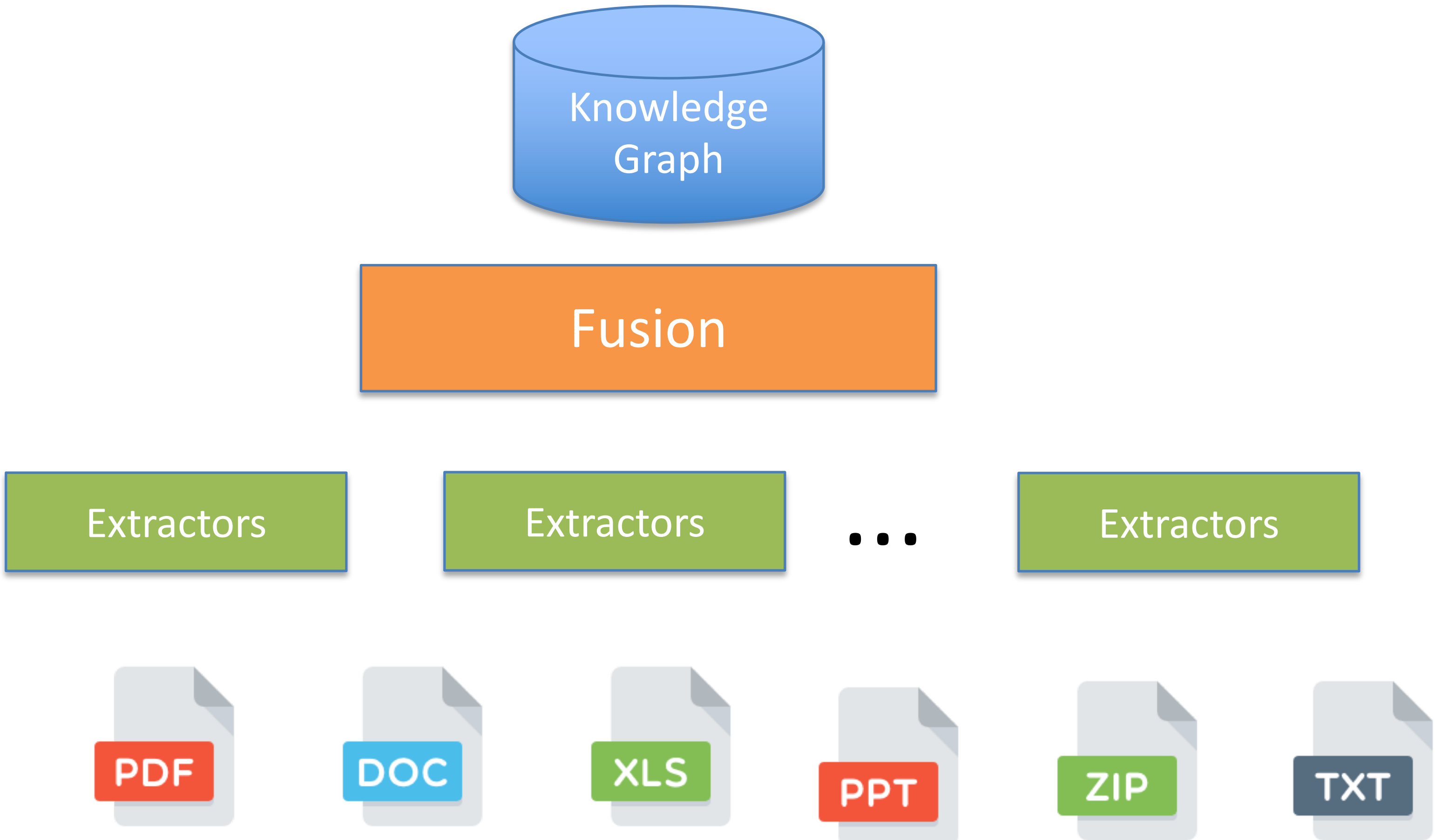
- Search for related evidence from data sources / knowledgebase
- Evaluate and aggregate the evidence and determine the correctness

- Assumption:
  - No errors due to extractors
  - Fact checking can be done using available data

DB Based Fact Checking

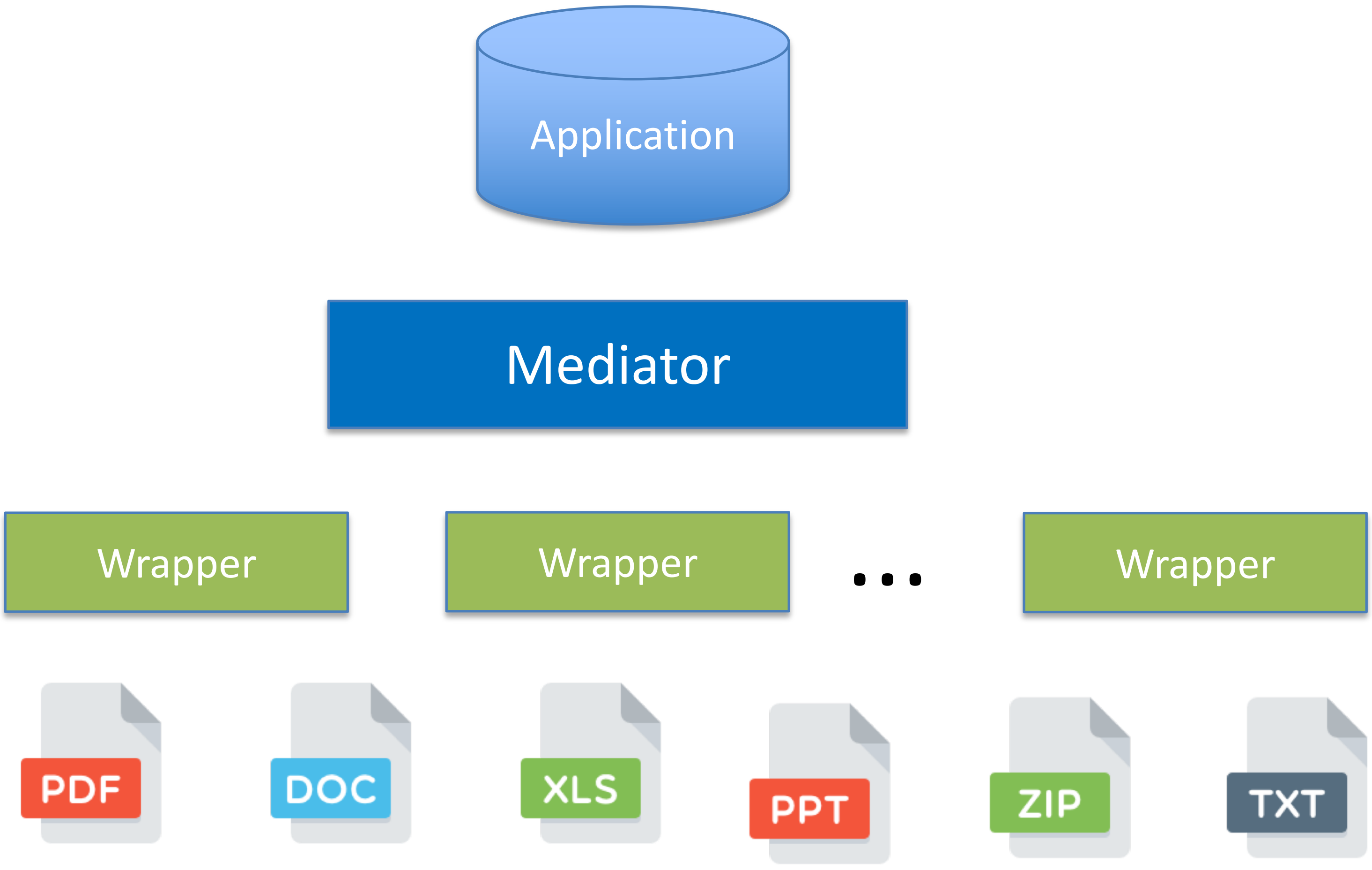
- Data Integration
- Data Fusion
- Crowdsourcing
- Knowledge Graphs

DB Based Fact Checking



[Dong, Gabrilovich, Heitz, et al. *VLDB*. 2014.]

DB Based Fact Checking



[Katsis, Papakonstantinou. *EDBS*. 2009]

Data Integration General

- DB community was one of the earliest to tackle discrepancy b/w data sources.
- Schematic discrepancy ➔ (Schematic) Data Integration.
- Inconsistency in data ➔ paraconsistent logics, data cleaning, etc.

(Schematic) Data Integration

- Mediated Schema as a view over each data source: *global as view (GAV)*.

- Each source as a view over mediated schema: *local as view (LAV)*.

- Hybrid: *GLAV*.

[Katsis, Papakonstantinou. *EDBS*. 2009]

Flights Example

- American Airlines:

Flight	DA	AA	DD	SDT	ADT	DG	SAT	AAT	AG
AA1007	TPA	MIA	12/01/2011	13:55	14:07	F78	15:00	14:57	D5

- Air Travel Center:

Flight	DA	AA	DD	DT	AT
AA1007	TPA	MIA	12/01/2011	14:06	14:51

- Orbitz:

Flight	DA	AA	DD	SDT	ADT	DG	SAT	AAT	AG
AA1007	TPA	MIA	12/01/2011	13:55	13:57	F78	15:00	14:57	D5

Flights Example (contd.)

- Consistent Query Answers?
- Logic(s) of Inconsistency?
- Metric FDs to the rescue?
- Which source is correct?
  - grade facts (claims) as well as sources (claimants).

Data Quality and Fake News

Source	Person	Institution
S1	Jiawei Han	UIUC
S2	J Han	University of Illinois at Urbana-Champaign
S3	Jiawei Han	SFU
S4	Jiawei Han	UCLA

Source	Person	Birthplace
S1	Barack Obama	Hawaii
S2	Barack Hussein Obama	Kenya
S3	Barack H. Obama	Kenya
S4	Barack Obama	Honolulu

Data Quality and Fake News

- Inconsistency as a web data quality issue
- Fake news is just another pernicious manifestation
- How can we leverage prior research for fact checking?

Truth Discovery

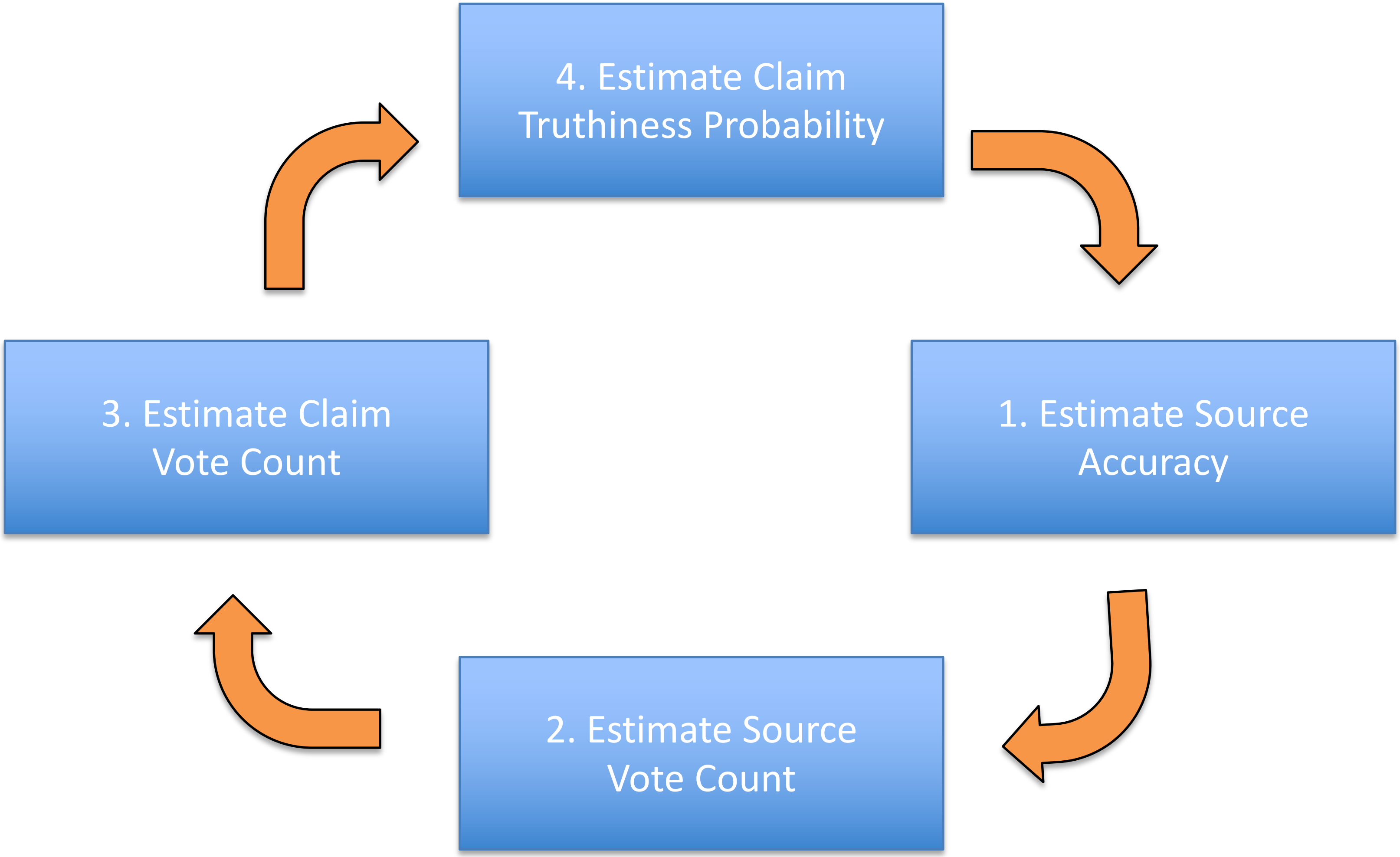
- Given:
  - set of sources
  - Claims made by sources
- Output:
  - For each claim, probability that it is true
- Related to computational problem P2
- Intuition:
  - Some sources are more trustworthy
  - Trustworthy sources are usually right

Prior Approaches

- EM like Approaches

- Supervised Approaches

EM like Approaches: ACCU



[Dong, Berti-Equille, Srivastava. *VLDB*. 2009]

Handling Correlated Sources

Correlation via Copying

- How to detect copying?
  - If two sources share a lot of false values, they are more likely to be dependent.

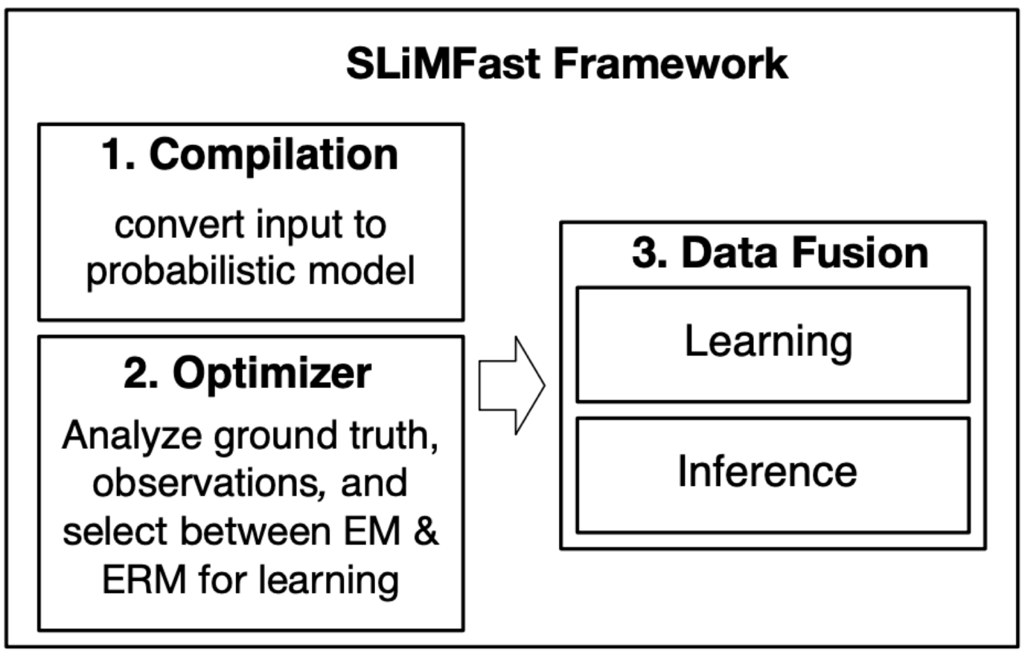
- How to determine who copied from whom?
  - If source S1 copied from S2, then accuracy of S1 on entire data will be different from accuracy of S1 on common data

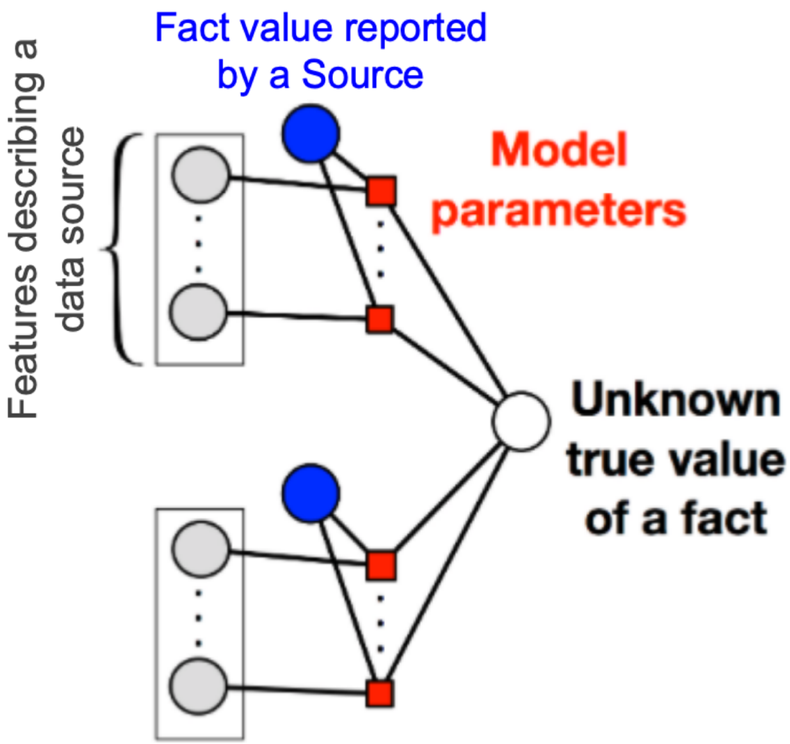
[Dong, Berti-Equille, Srivastava. *VLDB*. 2009]

Supervised Truth Discovery

- How can we use existing fact checkers?
- Idea: Leverage domain specific features to reduce labeled data
  - Age of news source
  - Content quality
  - Number of articles, topics, visitors
  - Source and topic partisanship

SLiMFast





[Rekatsinas, Joglekar, Garcia-Molina et al. *SIGMOD*. 2017.]

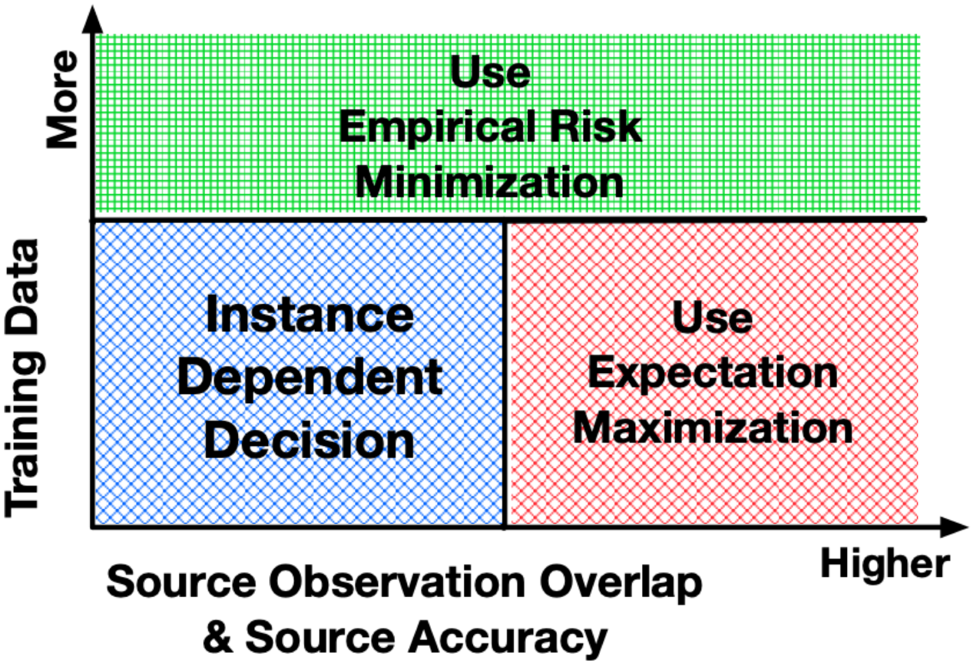
SLiMFast

Supervised Learning:

- lot of training data

Unsupervised Learning:

- high average accuracy of data sources
- high density of source observations



[Rekatsinas, Joglekar, Garcia-Molina et al. *SIGMOD*. 2017.]

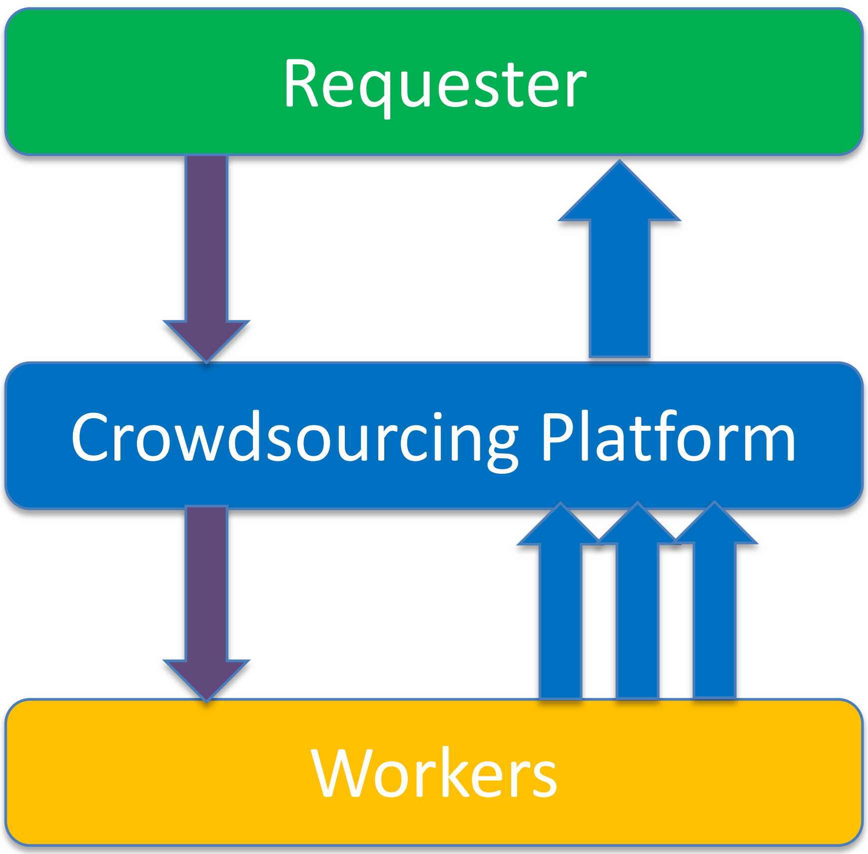
Fact Checking and Data Fusion

- Goal is to aggregate conflicting data sources
- Relies on estimating data source reliability
- Intuition: Reliable data sources have typically accurate results
- Data source quality and true labels are often unknown
- Correlation occurs due to copying/partisanship

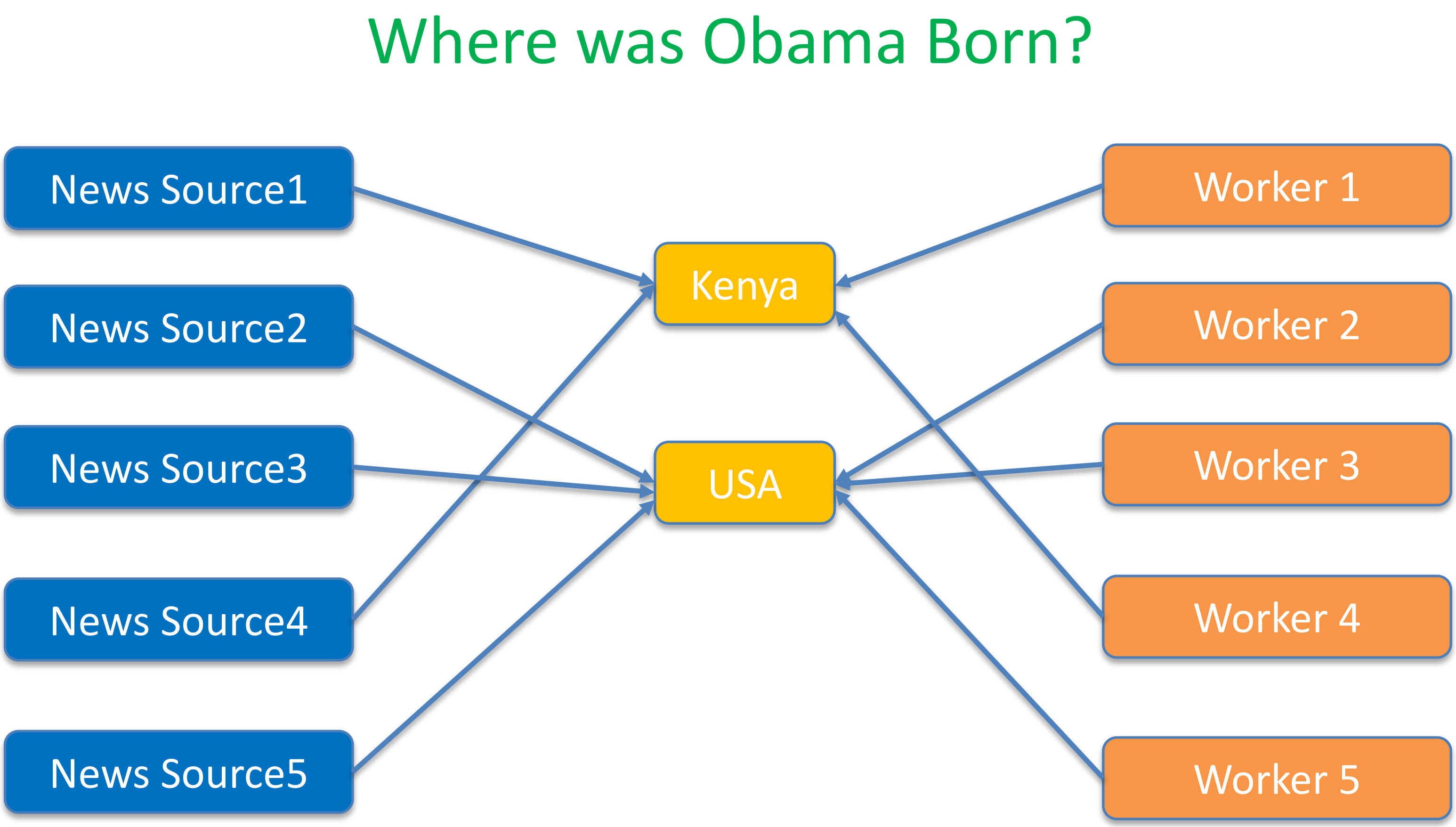
[Gao, Li, Zhao et al. *PVLDB*. 2015]

Crowdsourcing

Crowdsourcing Workflow



Crowdsourcing and Data Fusion



Crowdsourcing and Data Fusion

Data Fusion / Fact Checking	Crowdsourcing
Aggregate conflicting data sources and claims	Aggregate conflicting worker answers
Estimate data source reliability	Estimate worker quality
Reliable sources typically accurate results	Reliable workers typically produce accurate responses
Source quality and true labels are unknown	Worker quality and true labels are unknown

[Gao, Li, Zhao et al. *PVLDB*. 2015]

4c. Detection – Database Approaches: Knowledge Graphs

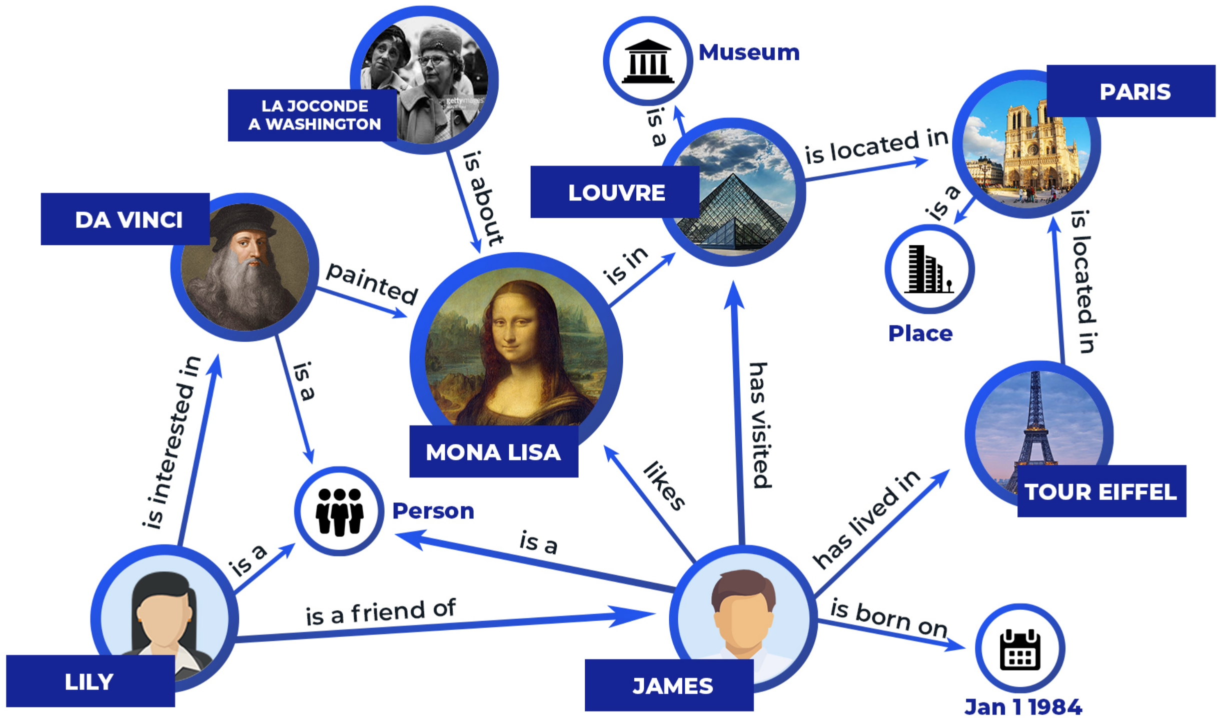
Knowledge Graph Construction

Sources include:

- News copora
- Wikipedia entries
- Web tables
- Databases

Knowledge Graph Representation

KG's represent *known facts* as a set of SPO triples of the form: **(Subject, Predicate, Object)**



Assumption. KG stores facts that are collected from *trusted* sources

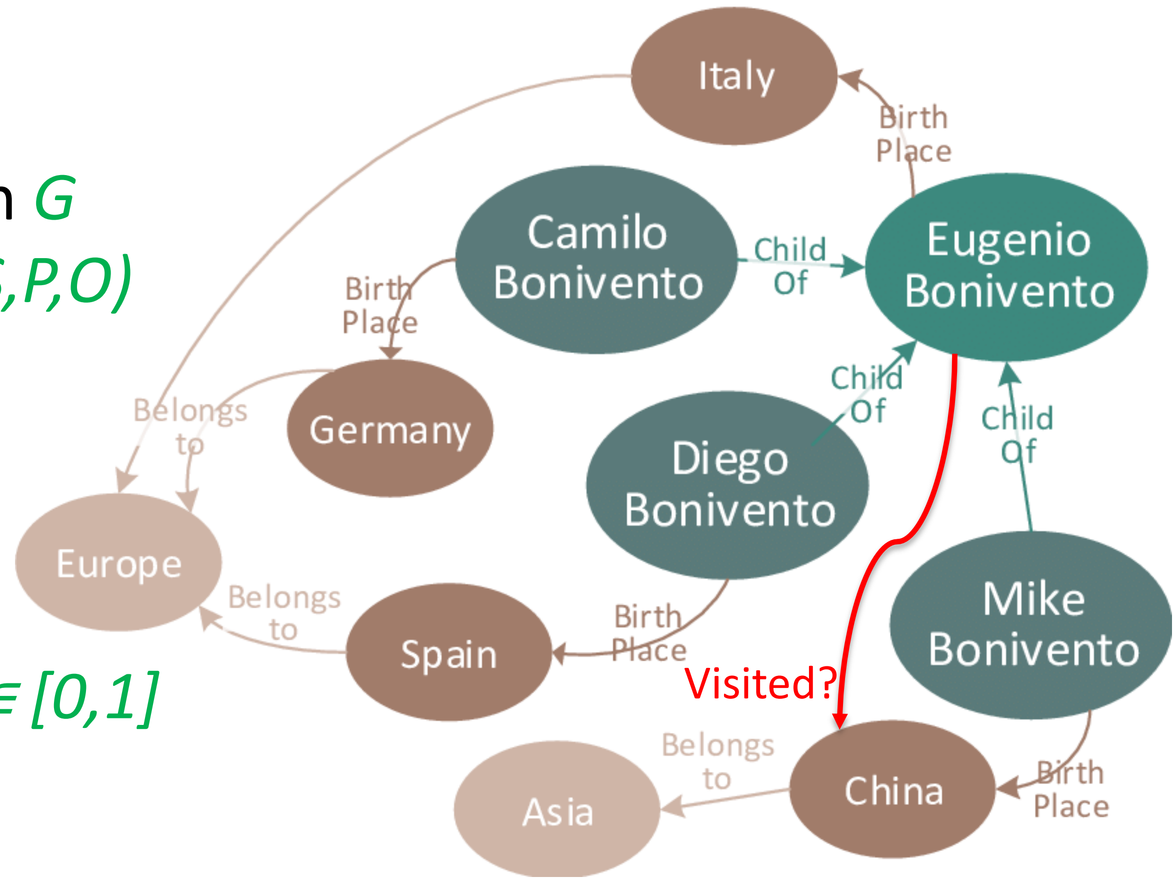
Fact Checking with KG's

Simple fact checking can be cast as **Triple Verification (P4)**.

Input:

Knowledge Graph *G*

claim triple *C = (S,P,O)*



Output:

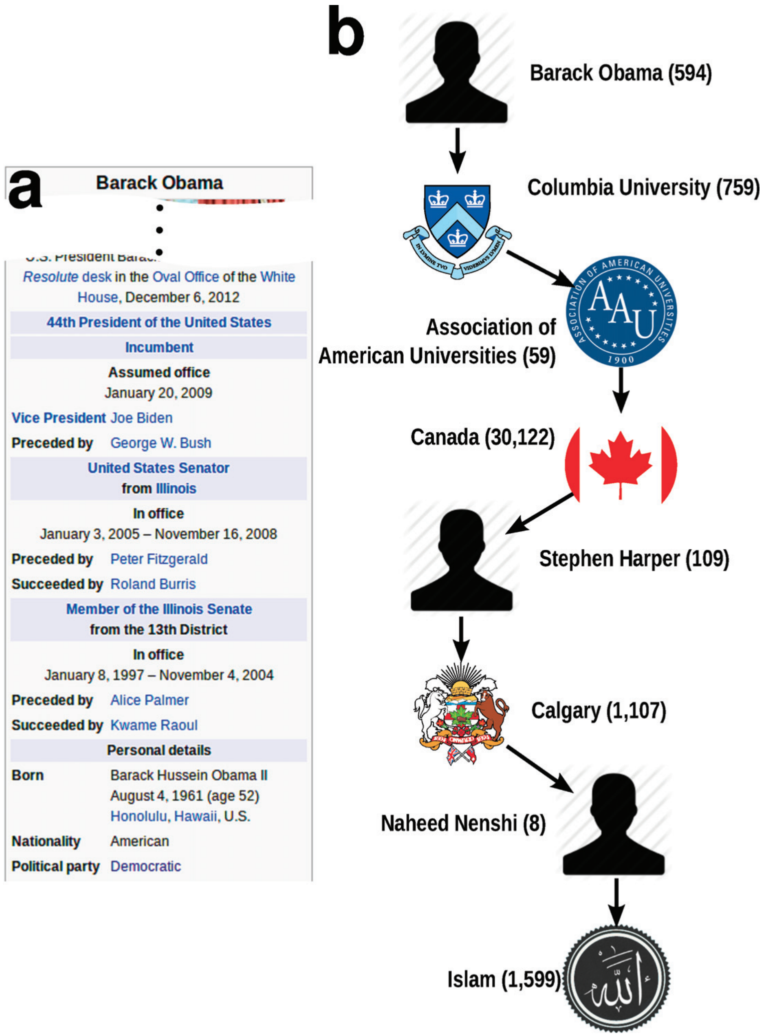
Truth Score $\tau(C) \in [0,1]$

[Image: Morales et al. *ICWE* 2017.]

Fact Checking with KG's

Challenges:

1. KG's are often *incomplete*
2. Claims may contain *long-range dependencies*
3. Finding *scalable* solutions
4. Providing *explanations*

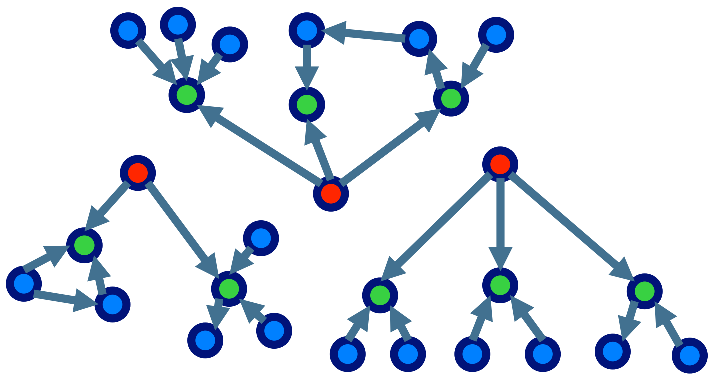


Approaches: Similarity

Idea. Structural characteristics of *S* and *O* in KG is a proxy for similarity --> better truth score τ(C).

Frameworks. Katz centrality (1953), SimRank (2002), Local Path Index (2009), Path Entropy (2016)

Features. Degree, (shortest) paths, neighbourhood structure, etc.



Does **not** leverage node/edge labels or types!

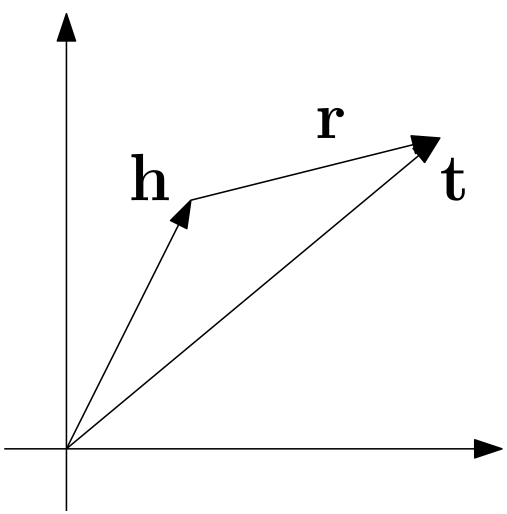
Fast, but relatively inaccurate.

[Jeh & Widom. *KDD* 2002.]

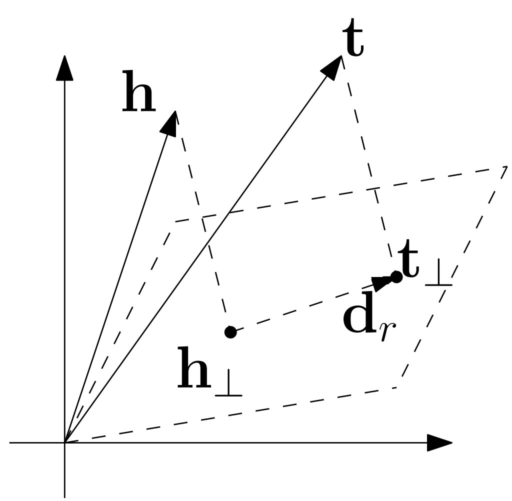
Approaches: Vector Space

Idea. Embed entities and relations in low-dimensional vector space and do link prediction.

Frameworks. TransE (2013), TransH (2014), TransR (2015), DistMult (2015), ProjE (2017), SimplE (2018)



(a) TransE



(b) TransH

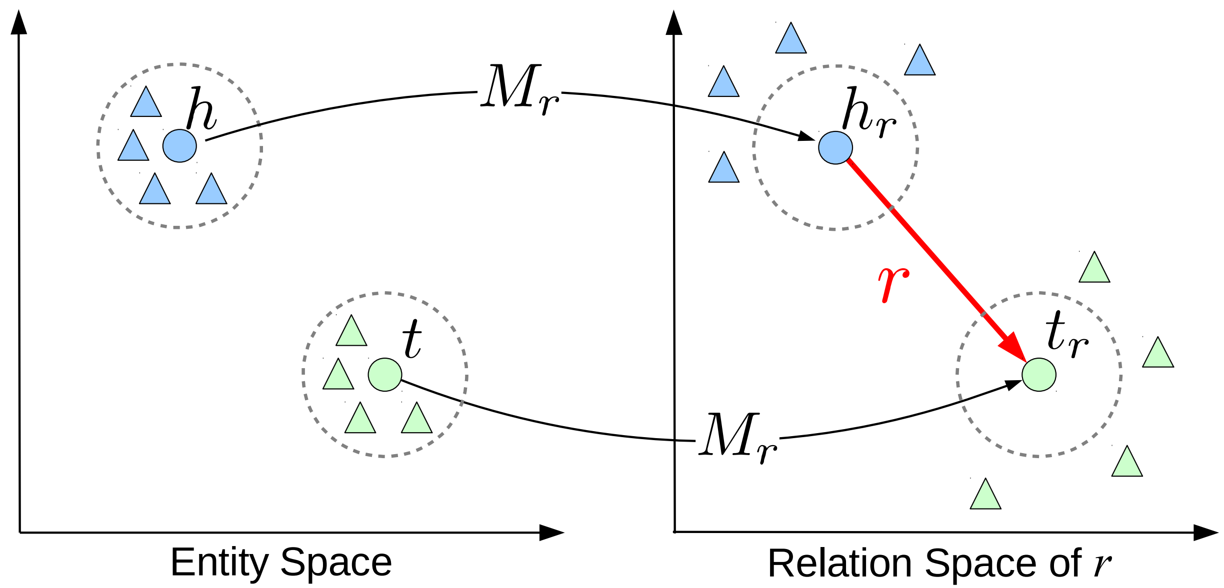
Translation in relation-specific hyperplane

[Bordes et al. *NIPS* 2013.]

[Wang et al. *AAAI* 2014.]

Approaches: Vector Space

Frameworks. TransE (2013), TransH (2014), TransR (2015), DistMult (2015), ProjE (2017), SimplE (2018)



Leverage separate entity and relation spaces

[Lin et al. *AAAI* 2015.]

Approaches: Vector Space

Vector space approaches are achieving higher accuracy as the models become more sophisticated, but…

Limitations:

- most lack interpretable evidence
- suffers from inverse relation bias

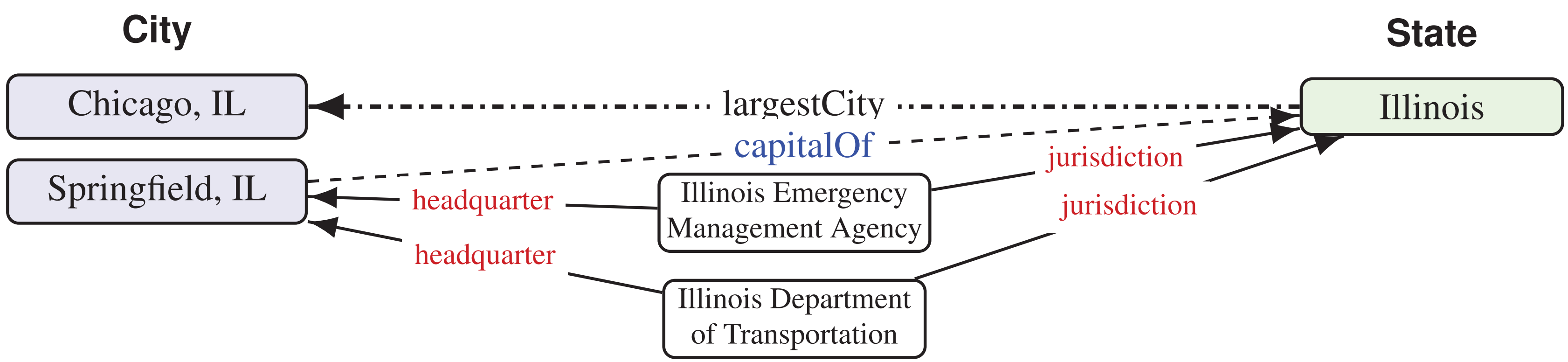
	FB15k						FB15k-237					
	Raw			Filtered			Raw			Filtered		
Model	MR↓	Hits@10↑	MRR↑	FMR↓	FHits@10↑	FMRR↑	MR↓	Hits@10↑	MRR↑	FMR↓	FHits@10↑	FMRR↑
TransE [3]	243.0 201.0	34.9 43.4	— 18.44	125.0 70.2	47.1 61.8	- 30.7	- 440.2	- 29.8	- 11.9	- 250.8	- 42.5	- 18.0
TransH [16]	211.0 213.8	42.5 47.3	- 28.3	84.0 69.3	58.5 70.1	- 16.3	- 511.8	- 29.0	- 10.5	- 309.8	- 42.9	- 16.3
TransR [9]	226.0 236.4	43.8 47.2	- 16.2	78.0 82.7	65.5 71.9	- 29.7	- 544.9	- 27.9	- 9.9	- 337.0	- 42.9	- 16.2
TransD [7]	211.0 209.8	49.4 47.4	- 16.3	67.0 65.4	74.2 70.4	- 28.3	- 506.9	- 29.4	- 10.4	- 305.2	- 42.8	- 16.2

[Akrami et al. *CIKM* 2018.]

Approaches: Rule Mining

Idea. Interpret graph patterns or paths as "rules" with matches in KG providing evidence for *C*.

Supervised: PRA (2010), PredPath (2016), Gfact (2018)
Unsupervised: KL (2015), KL-REL (2017), KS (2017)



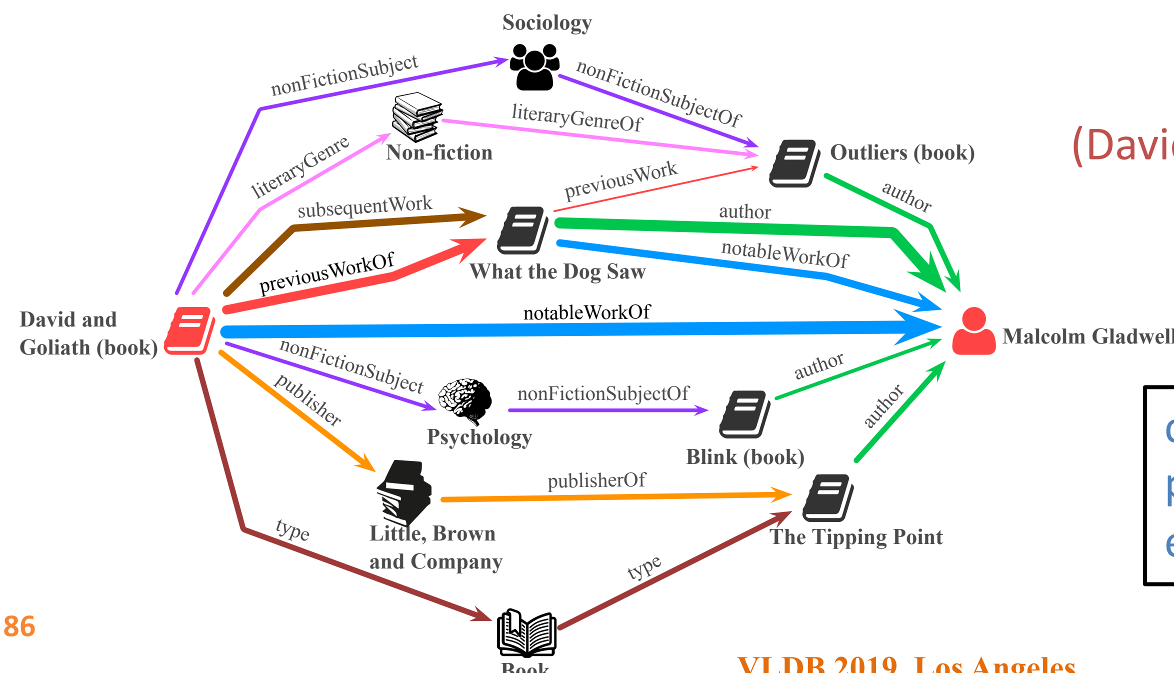
Mines patterns that *uniquely* define "capitolOf" relation

[Shi et al. *Knowledge-Based Systems* 2016.]

Approaches: Rule Mining

Idea. Interpret graph patterns or paths as "rules" with matches in KG providing evidence for *C*.

Supervised: PRA (2010), PredPath (2016), Gfact (2018)
Unsupervised: KL (2015), KL-REL (2017), KS (2017)



(David & Goliath, author?, Malcom Gladwell)

collection of paths with "similar" predicates provide strong evidence for author relationship

[Shiralkar et al. *ICDM* 2017.]

Approaches: Rule Mining

Claim. Rule Mining approaches are accurate & interpretable.

CapitalOf #1	{city}	$\langle$headquarter^{-1}, jurisdiction$\rangle$ $\langle$location^{-1}, jurisdiction$\rangle$	{state}
CapitalOf #2	{city}	$\langle$location^{-1}, location$\rangle$ $\langle$isPartOf$\rangle$	{state}
Company CEO	{person}	$\langle$employer$\rangle$	{company}
US Civil War	{person}	$\langle$notable commander^{-1}, takePartIn$\rangle$	{battle}
NYT Bestseller	{person}	$\langle$notable work, previous work$\rangle$ $\langle$notable work, subsequent work$\rangle$	{book}
US President	{vice president}	$\langle$successor, president^{-1}$\rangle$	{president}

CEO	(parentCompanyOf, keyPerson)	32	News Corporation $\xrightarrow{\text{parentCompanyOf}}$ Sky TV plc $\xrightarrow{\text{keyPerson}}$ Rupert Murdoch
	(employerOf)	24	Twitter $\xrightarrow{\text{employerOf}}$ Dick Costolo
	(foundedBy)	24	Foxconn $\xrightarrow{\text{foundedBy}}$ Terry Gou
	(subsidiary, keyPerson)	20	Samsung $\xrightarrow{\text{subsidiary}}$ Samsung Electronics $\xrightarrow{\text{keyPerson}}$ Lee Kun-hee

Top patterns discovered by PredPath (top) & KnowledgeStream (bottom)

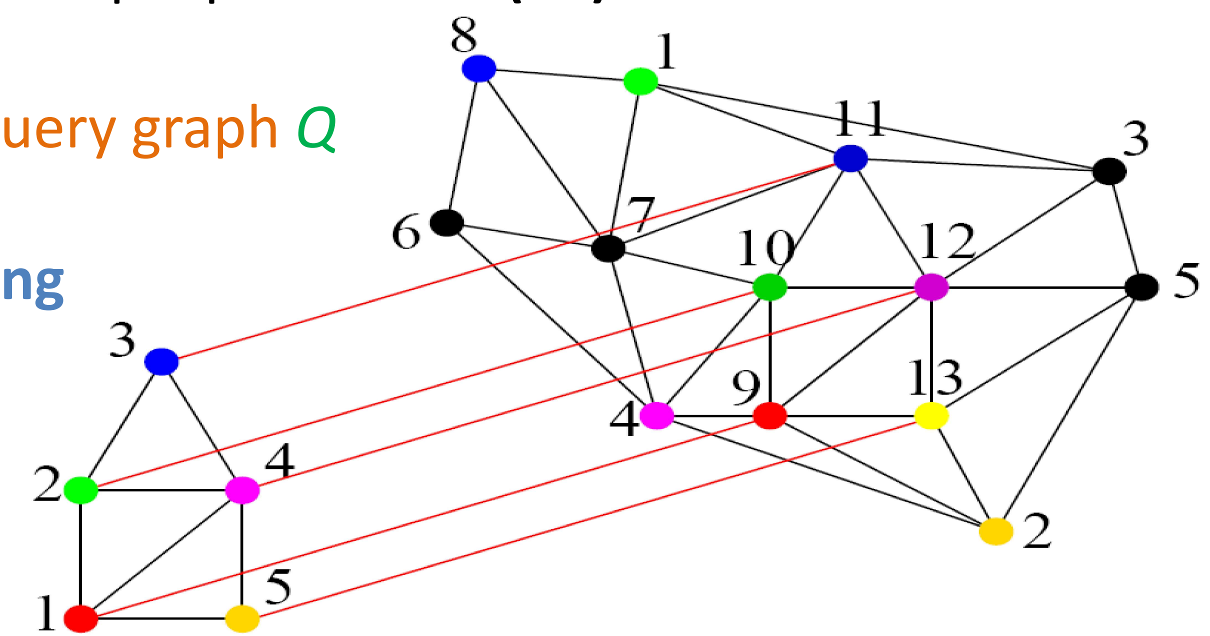
Beyond Link Prediction

How can we handle more complex claims?

E.g. claims involving many entities and multiple links connecting them with (possibly) unique predicates. **(P5)**

Matching against a query graph *Q*

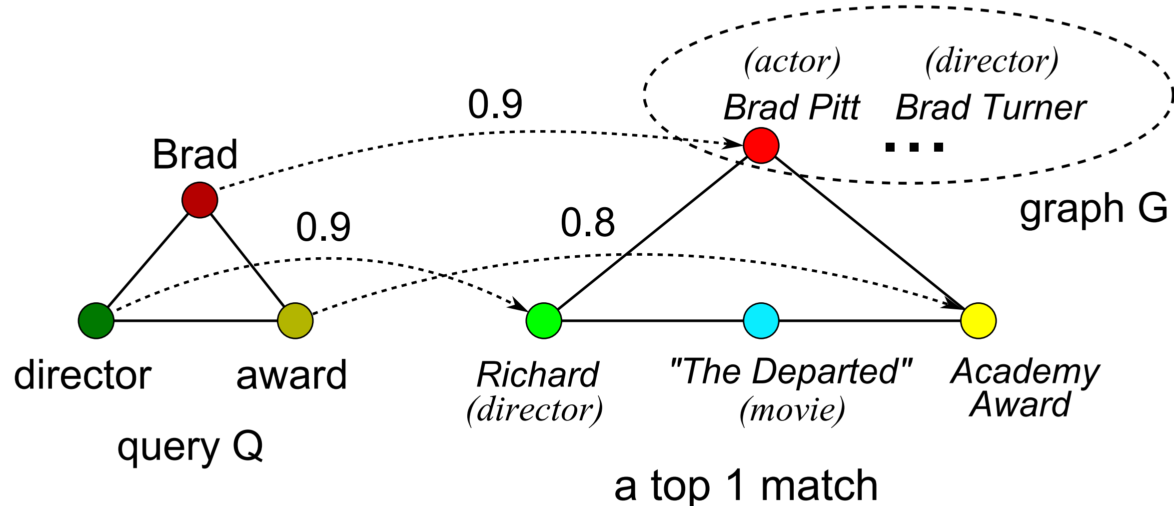
--> **Subgraph Matching**



Beyond Link Prediction

Frameworks. mtree (2013), SLQ (2014), Topk-EN (2015), STAR (2016)

Exhaustive SM is slow!



Overcome with TA-style approaches for answering top-k queries.

Pros. Complex queries and approximate node matching

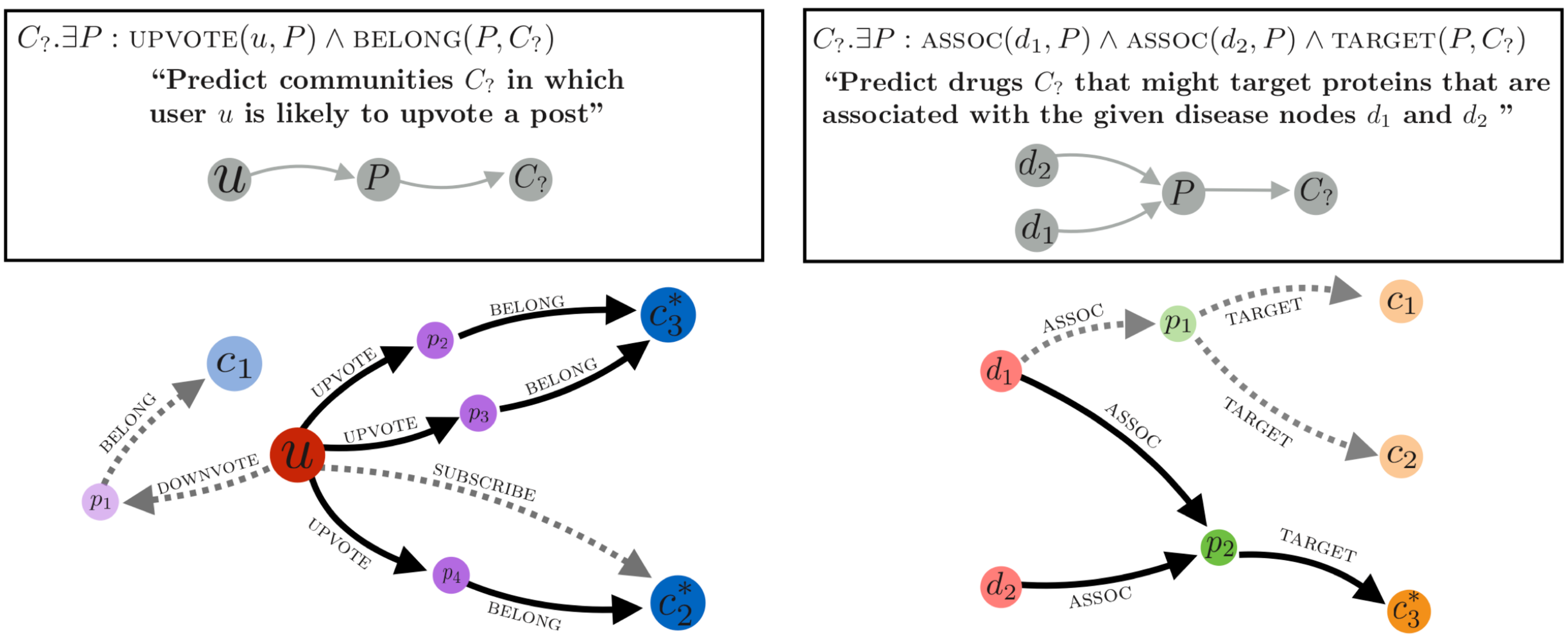
Cons. - Connectivity constraints do not consider edge predicate.
- Bounded path lengths.
- Restricted *Q* structure

[Yang et al. *ICDE* 2016.]

Beyond Link Prediction

Vector space approaches for conjunctive logical queries.

Idea. Embed nodes in low-dimensional space and represent logical operators as learned geometric operations in embedding space.



Valid queries form a DAG w/ *anchors* as sources and *targets* as the unique sink

[Hamilton et al. *NIPS* 2018.]

Beyond Link Prediction

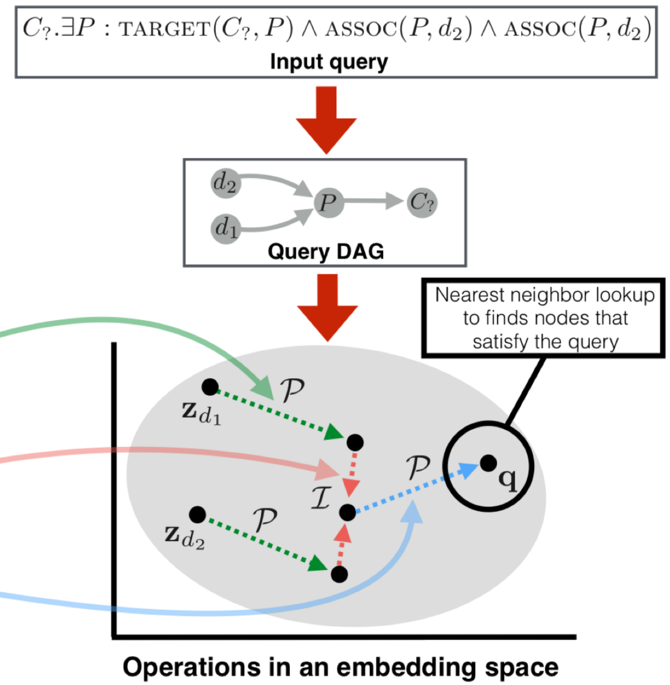
Idea. Embed nodes in low-dimensional space and represent logical operators as learned geometric operations in embedding space.

Algorithm 1: Query embedding generation

Input : Query anchor nodes $\mathcal{A}$, query variable nodes $\mathcal{B}$, query edges $\mathcal{E}_q$, a map d_q from query variables to their degree in the query DAG

Output : Query embedding $\mathbf{q}$

Q = dictionary mapping every $\mathcal{V}_i \in \mathcal{B}$ to an empty set;
for $\tau(v_i, V_j) \in \mathcal{E}_q \ : \ v_i \in \mathcal{A}$ **do**
  $Q[V_j] = Q[V_j] \cup \mathcal{P}(\mathbf{z}_{v_i}, \tau)$
while $|Q.key\_set| > 0$ **do**
  A = empty dictionary;
  for $V_i \in Q.key\_set \ : \ |Q[V_i]| = d_q(V_i)$ **do**
    $A[V_i] = \mathcal{I}(Q[V_i])$;
    delete $Q[V_i]$;
  for $V_i \in A.key\_set$ **do**
    for $\tau(V_j, V_k) \in \mathcal{E}_q \ : \ V_j = V_i$ **do**
      $Q[V_k] = Q[V_k] \cup \mathcal{P}(A[V_i], \tau)$;
return $A[V_?]$;



Projection: translates in direction determined by edge type.
Intersection: set intersection in embedding space on node embeddings of the same type.

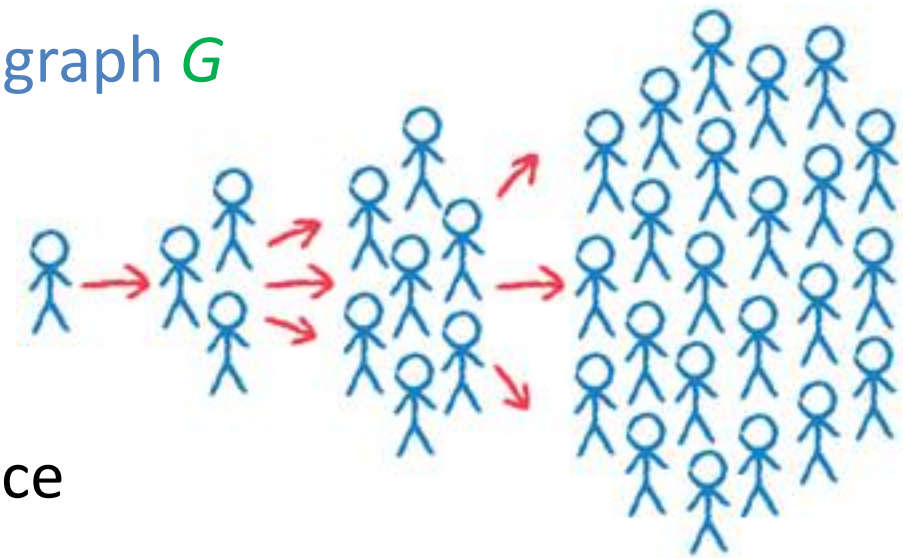
5. Mitigation & Intervention

Mitigation & Intervention: Influence Maximization Models

Influence Maximization

Model a social network as a graph *G*

- Edges = relationships
- Nodes = users



Edge weights are estimates for the probability of influence

IM Problem. Achieve widespread adoption of a product by initially “seeding” a few users.

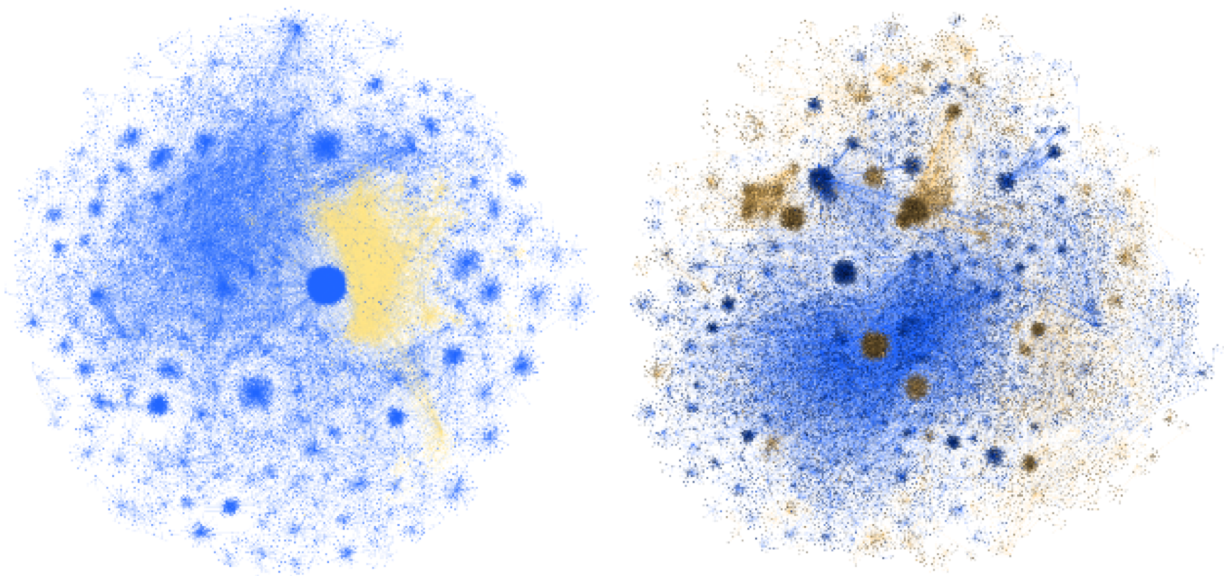
Idea. Influential users trigger a cascade of influence

[Kempe et al. *KDD* 2003.]

Mitigation via Truth Campaigns

Idea. Combat fake news with a truth campaign **(P6)**.

Goal. Disseminate the truth such that the number of users who end up adopting the fake news is minimized. (**NP**-hard)



blue = truth
yellow = fake

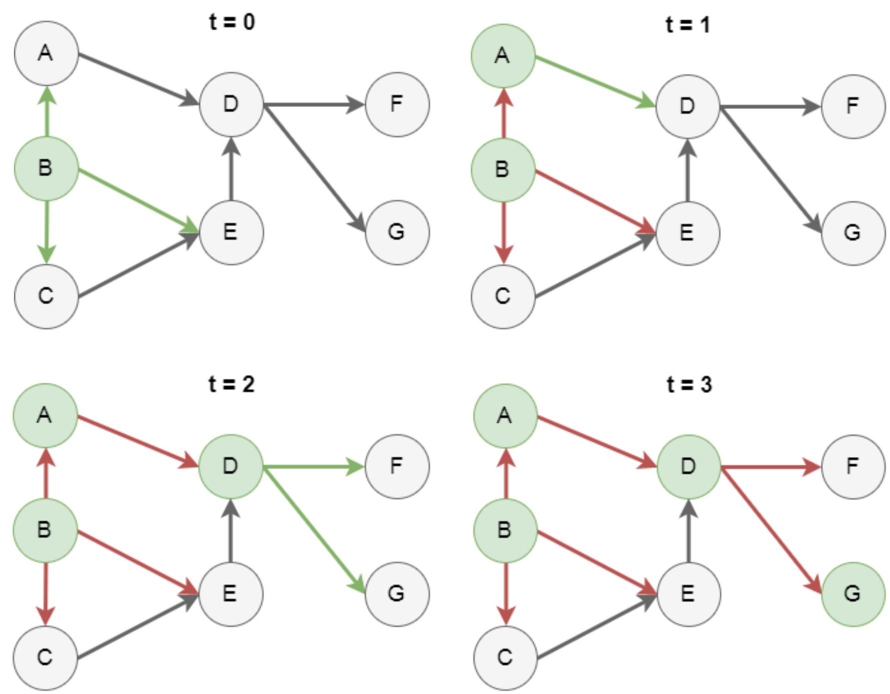
adoption is *exclusive!*

[Image: http://cnets.indiana.edu/blog/2016/12/21/hoaxy/]

Independent Cascade Model

Diffusion of information under IC occurs in a series of rounds:

1. Activate seed set
2. in each round, newly active nodes have single chance to activate inactive neighbours
3. Use influence probabilities on edges to resolve activations
4. Active nodes do not de-activate



seed = *B*

[Image: Gursoy & Durahim. *arXiv.* 2018.]

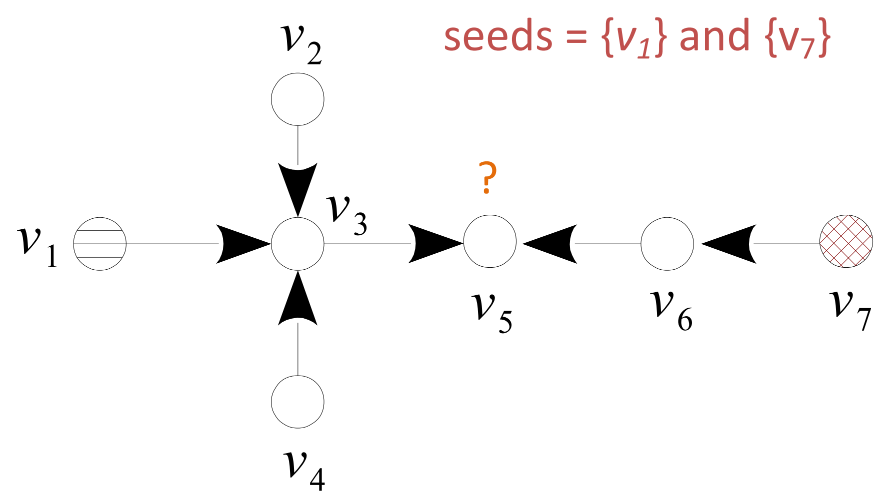
[Kempe et al. *KDD* 2003.]

Competitive IC Model

Diffusion under CIC has additional considerations:

1. Activates two seed sets
2. Cascades can share edge liveness or propagate separately
3. Active nodes do not switch campaigns
4. Competition requires defining a tie-breaking rule (e.g. positive/negative dominance, proportional probability, etc.)

seeds = {v_1} and {v_7}



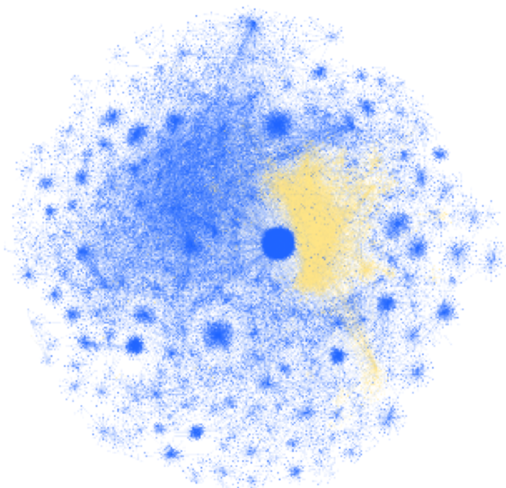
[Bharathi et al. *WINE* 2007.]

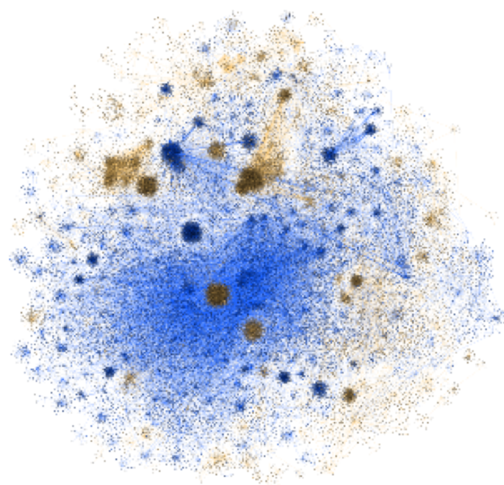
[Budak et al. *WWW* 2011.]

Mitigation via Truth Campaigns

Objective. Select *k* seeds for truth campaign that maximizes number of users prevented from adopting the misinformation.

Solution. When campaigns share possible worlds (i.e. edge liveness shared) then objective is monotone & submodular --> Greedy yields *(1 – 1/e)*-approximate solution.



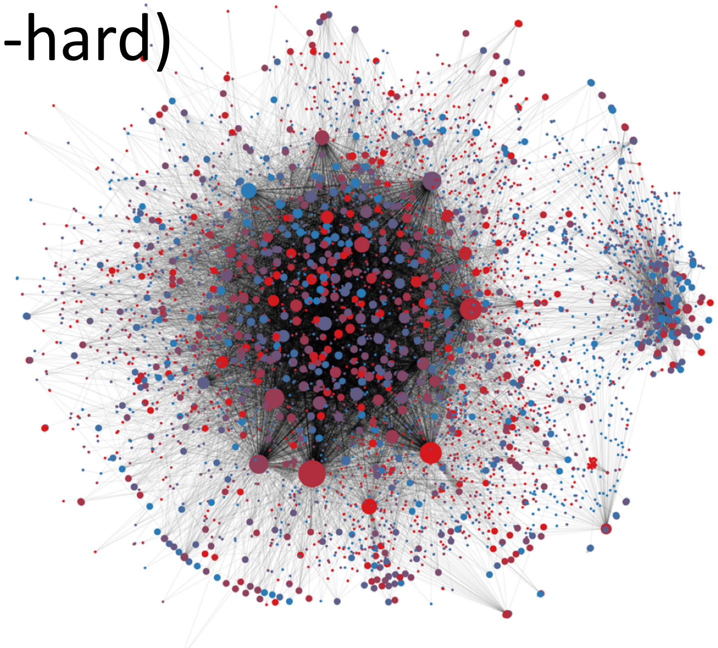


[Budak et al. *WWW* 2011.]

[Tong et al. *TNSE* 2017.]

Mitigation via Truth Campaigns

Alternative Goal. Select the minimum number of nodes to seed in the truth campaign to protect at least a *β* fraction of the network. (**NP**-hard)



Solution. Greedy selection of *β-Node Protectors* returns set of size at most $|OPT| + O(1/e * \beta N)$.

[Nguyen et al. *WebSci* 2012.]

Hard Intervention Techniques

Question. What if a truth campaign isn't effective enough?

- Consider network modification via edge removal. **(P7)**

Problem. Select *k* edges to remove from *G* such that the number of users adopting the fake news is minimized.

C1. *Total* number of edges (cardinality)

C2. limit edges that can be removed from *each* node (matroid).

Credit Distribution Model:

C1+2 APX-hard --> monotone submodular maximization

Linear Threshold:

C1 --> monotone supermodular minimization

[Khalil et al. *KDD* 2014.]

[Medya et al. *arXiv.* 2019.]

Mitigation & Intervention: Epidemiological Models

Virus Propagation Models

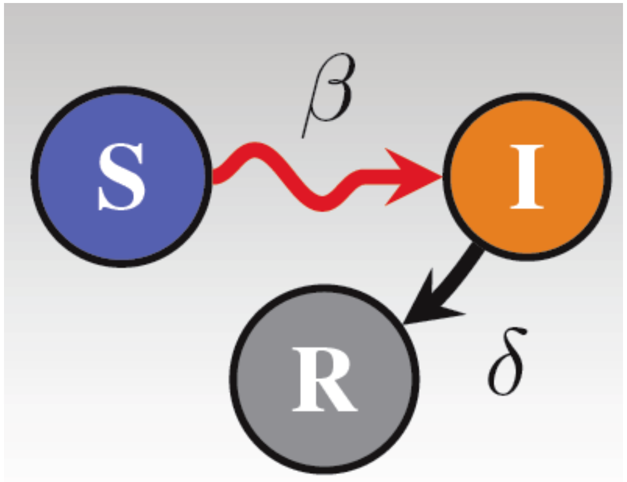
Given a graph *G* where edges represent contact relationships and nodes represent users.

VPM's defined by states and corresponding transitions.

For *SIR*, each node is in one of three states:

1. Susceptible (i.e. healthy)
2. Infected
3. Recovered (can't be re-infected)

Other VPMs: *SIS*, *SEIR*, *SIHR*, *SEIZ*



[Bettencourt et al. *PHYSICA A* 2006.]

Intervention via Immunization

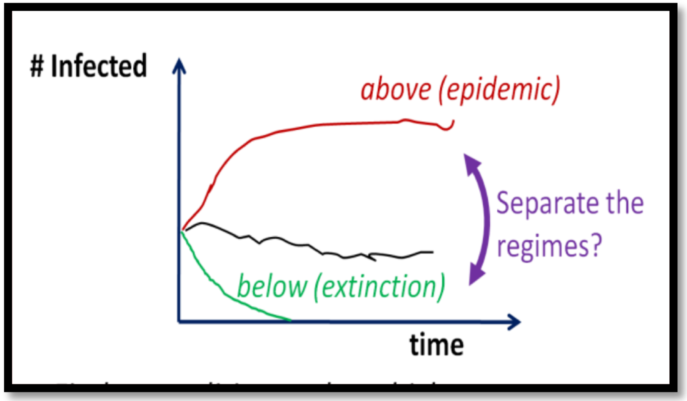
Goal. Choose best nodes/edges to remove (immunize). **(P7)**

Two settings:

- Pre-emptive: choose nodes to remove *before* epidemic starts
- Reactive: immunization occurs *after* epidemic starts

Prior work on VPM's studies the *epidemic threshold (ET)* which determines if a virus will die out or not.

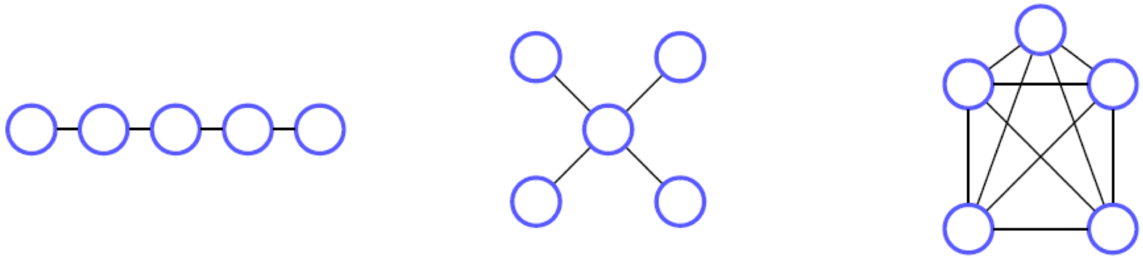
Result. ET depends only on first eigenvalue λ of adj matrix and a VPM dependent constant



[Prakash et al. *ICDM* 2011.]

Pre-emptive Immunization

Observe. Increasing λ --> increasing vulnerability of network



(a)Chain(λ_1 = 1.73) **(b)Star(λ_1 = 2)** **(c)Clique(λ_1 = 4)**

Goal. Select nodes that maximize the decrease of λ.

Solution. Approximate "eigen-drop" via matrix perturbation theory. Resulting objective is monotone & submodular.

[Tong et al. *ICDM* 2010.] [Prakash et al. *SDM* 2013.]

Reactive Immunization

Reactive: immunization occurs after epidemic starts.

Observe. The reactive immunization problem is a *special case* of the general CIC-based mitigation problem:

- Virus = misinformation & inoculation = truth
- Truth is static (i.e. edge probabilities are all zero)
- Thus, **NP**-hard and not submodular!

Proposed Solution:

1. Simplify graph by merging infected nodes into "super node"
2. Design optimal algorithm for trees (DAVA-tree)
3. Construct dominator tree *T* from *G* --> run DAVA-tree on *T*

[Zhang & Prakash. *SDM* 2014.]

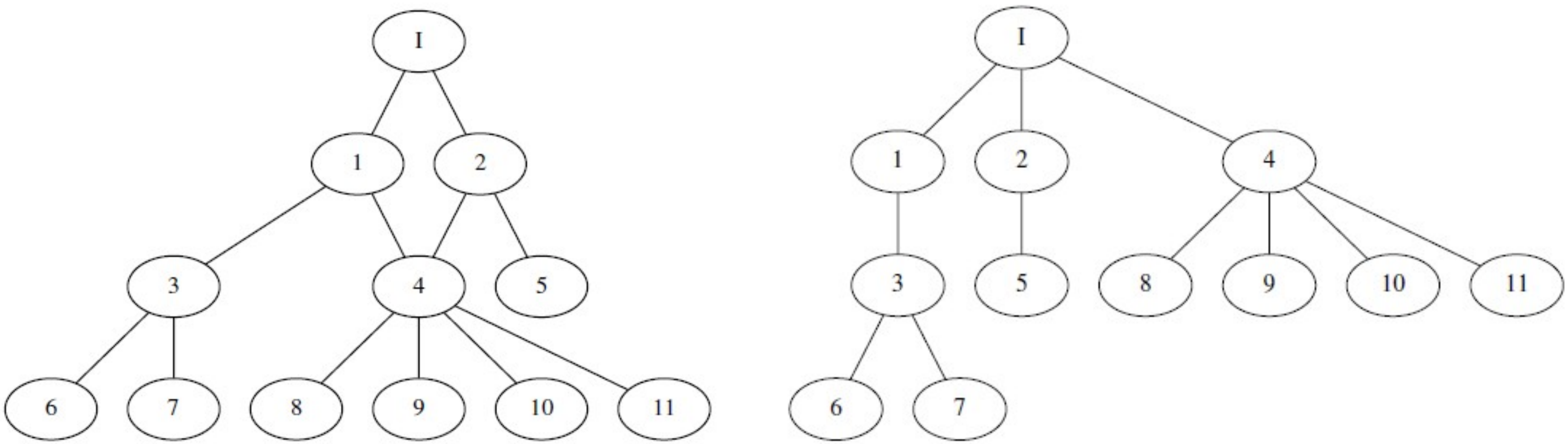
[Zhang & Prakash. *TKDD* 2015.]

Reactive Immunization

Construct dominator tree *T* from *G* --> run DAVA-tree on *T:*

u dominates *v*	⟺	every path from *I* to *v* contains *u*

(u,v)∃in *T* if *u* dominates *v* AND every other dominator of *v* dominates *u*



Merged Graph

Dominator Tree

Weighting *T* is **#P**-hard --> use maximum propagation path probability.

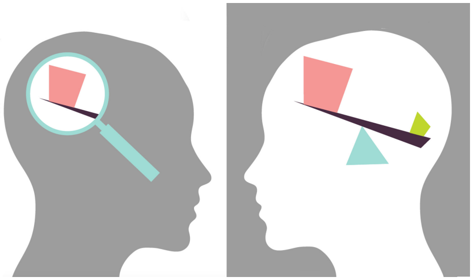
Mitigation & Intervention:
Soft Touch Techniques

Softer Touch Techniques

Recent attempts by major companies to combat fake news incorporate "gentle nudges" away from misinformation.

Question. What role does human decision making play in the adoption and propagation of misinformation and how can technology enable humans to make better decisions?

Informing users about different cognitive biases that humans are susceptible to can be leveraged in the design of intervention tactics.

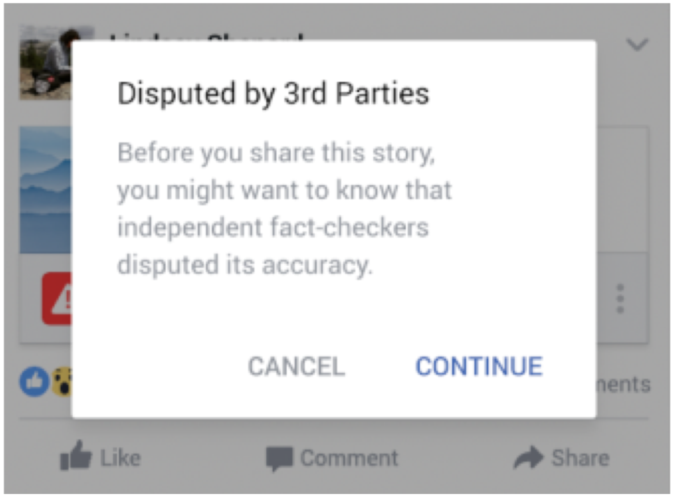


[Konstantinou et al. *Co-Inform Project.* 2019.]

Softer Touch Techniques

Facebook:

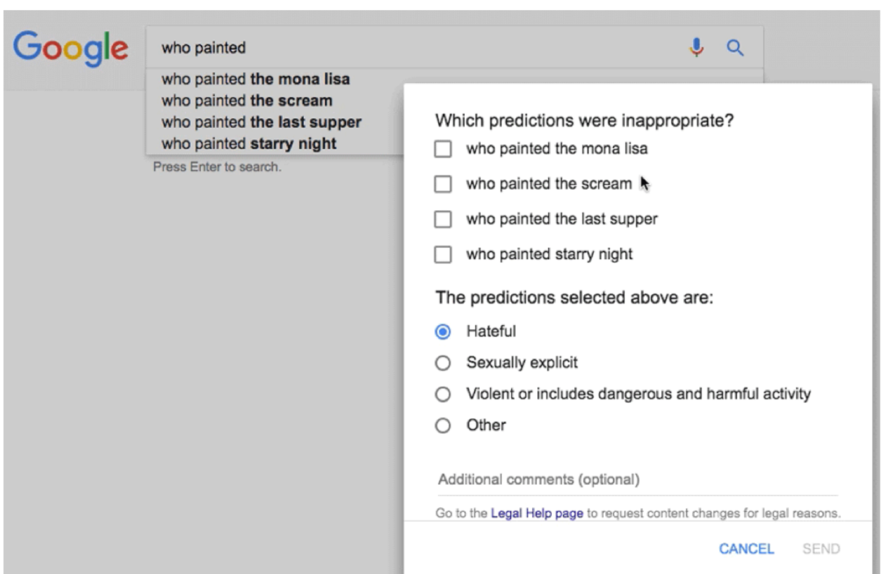
- Flagging stories as "disputed" by third-party fact checkers
- Disputed stories appear lower in News Feed
- Attempting to share a disputed story comes with a warning
- Informed sharing (when reading an article makes user less likely to share is used as signal for ranking)

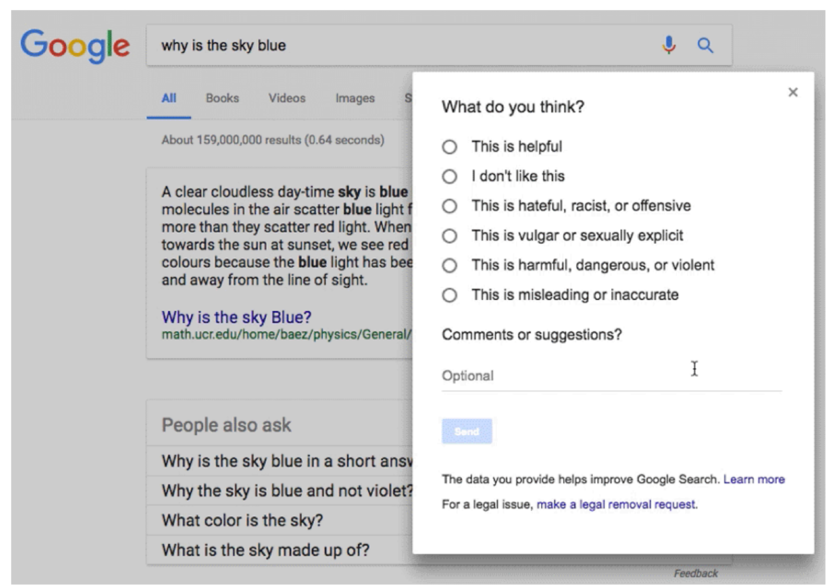


Softer Touch Techniques

Google:

- Improving search ranking via updated search quality guidelines for evaluators --> helps algorithms demote fake content in search results
- Easier ways to provide direct feedback on autocomplete predictions and featured snippets

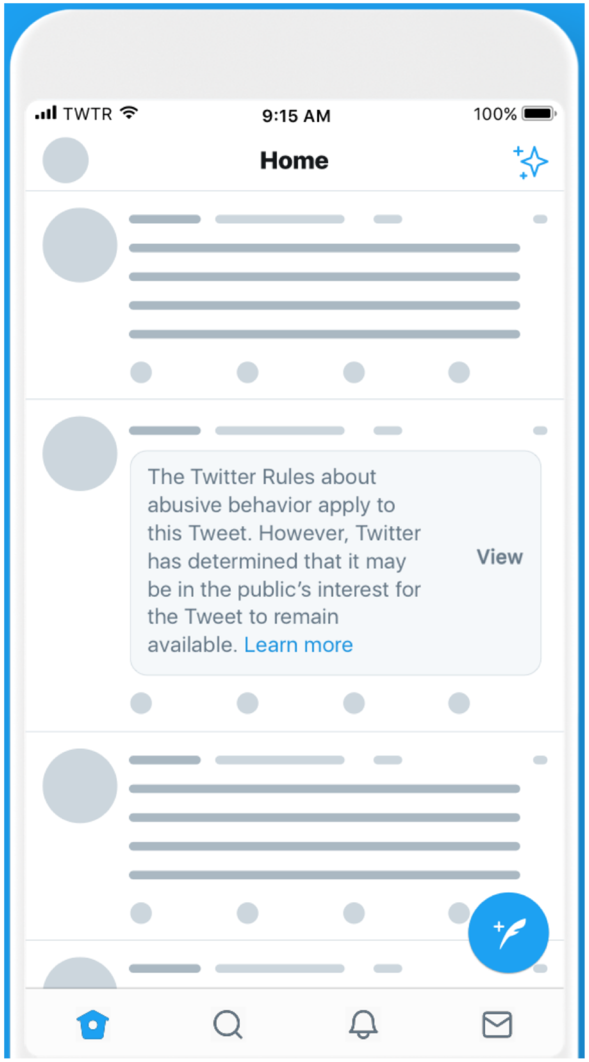




Softer Touch Techniques

Twitter:

- Notice providing additional clarity when posts that violate TOS are retained
- Applied to government/elected officials with >100K followers
- Must click through to see tweet
- Determination made by an interdisciplinary team (legal, policy, safety, etc.)
- Some content exempt and results in removal
- Tagged tweets are partially suppressed on platform

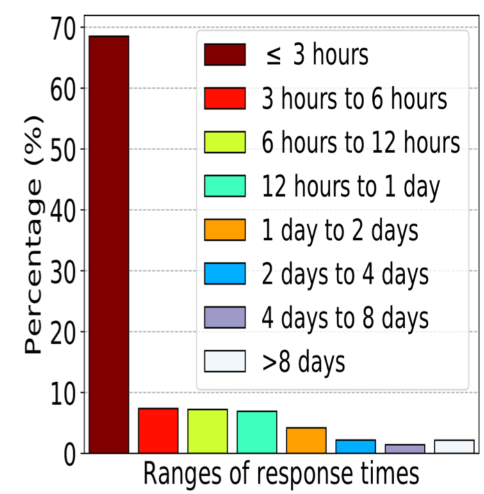


Softer Touch Techniques

Guardians are users who show interest in correcting false claims in online discussions by embedding URLs linking back to fact-checking sites.

- Majority of guardians post once or twice a year while a small subset are highly active (>200 posts).
- Verified accounts are more trustworthy and make up 2.2%
- Highly visible users (>5000 followers) make up 7.5%

1. Average response time is 2.26 days
2. 90% were posted within one day.



[Vo et al. *SIGIR* 2018.]

Softer Touch Techniques

Idea. Can we “outsource” the intervention task to guardians?

i.e. help guardians quickly access new interesting fact-checking URLs

Solution. Fact-checking URL recommender model that stimulates guardians to engage in intervention activities.

- Learn a model that recommends similar URLs to guardians whose interests are similar.
- Embedding based approach leverages URL content, network structure, and guardian post history.
- Outperforms SOTA approaches by 11-33%

5a. Fact Checking Ecosystem

Fact Checking Workflow

Monitor Sources

Spot / Extract Claims

Assess Claims

Report conclusion with supporting evidence

Fact Checking Approaches

Fact Checking Entities:

- Expert / Journalist
- Crowdsourcing through end users
- Human-Computer hybrid
- Fully automated

Expert based Fact Checking

FactCheck.org
A Project of the Annenberg Public Policy Center

TRUE
POLITIFACT
TRUTH-O-METER™

Snopes

Media Bias Fact Check

TruthOrFiction?
Seeking truth and exposing fiction since 1999

IFCN code of principles has 69 signatories so far

Expert based Fact Checking

Advantages

- Fact checking is often thorough
- Better credibility
- Can handle nuanced claims
- Can produce detailed evidence of fact checking

Disadvantages

- Not very scalable with average fact checking time of 7 days
- Harder to avoid human biases
- Not always easy to experts in esoteric domains

Crowdsourced Fact Checking

TRUTHSQUAD ON HEALTHCARE

Orrin Hatch, U.S. Senator

"87 million Americans will be forced out of their coverage under new health care regulations from President Obama."

Fact-check this quote:

Is this true or false?

True False Not Sure

WikiTRIBUNE

Evidence-based Journalism

Fiskkit

A better way to discuss the news

Paste article link here

TextThresher

Very few thriving projects!

What is a good hybrid workflow of users, experts and AI?

Crowdsourced Fact Checking

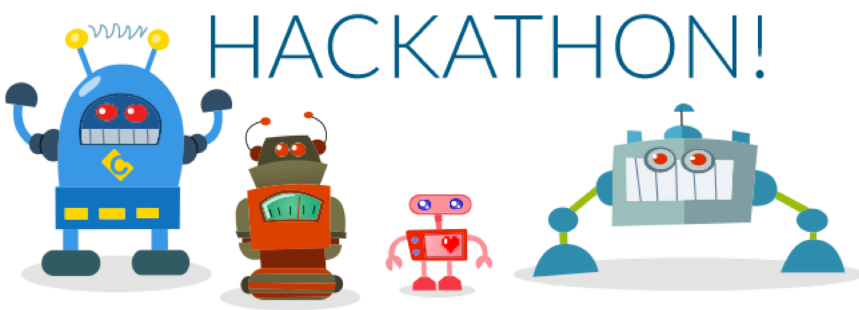
Advantages

- Leverage large number of users in a social media
- High scalability
- Easy to create workflows based on expertise and interest

Disadvantages

- Lower credibility
- Management of users is much harder
- Risk of manipulation by partisans
- Need to be aware of human biases
- Imbalance in volunteers for fact checking on popular vs important topics

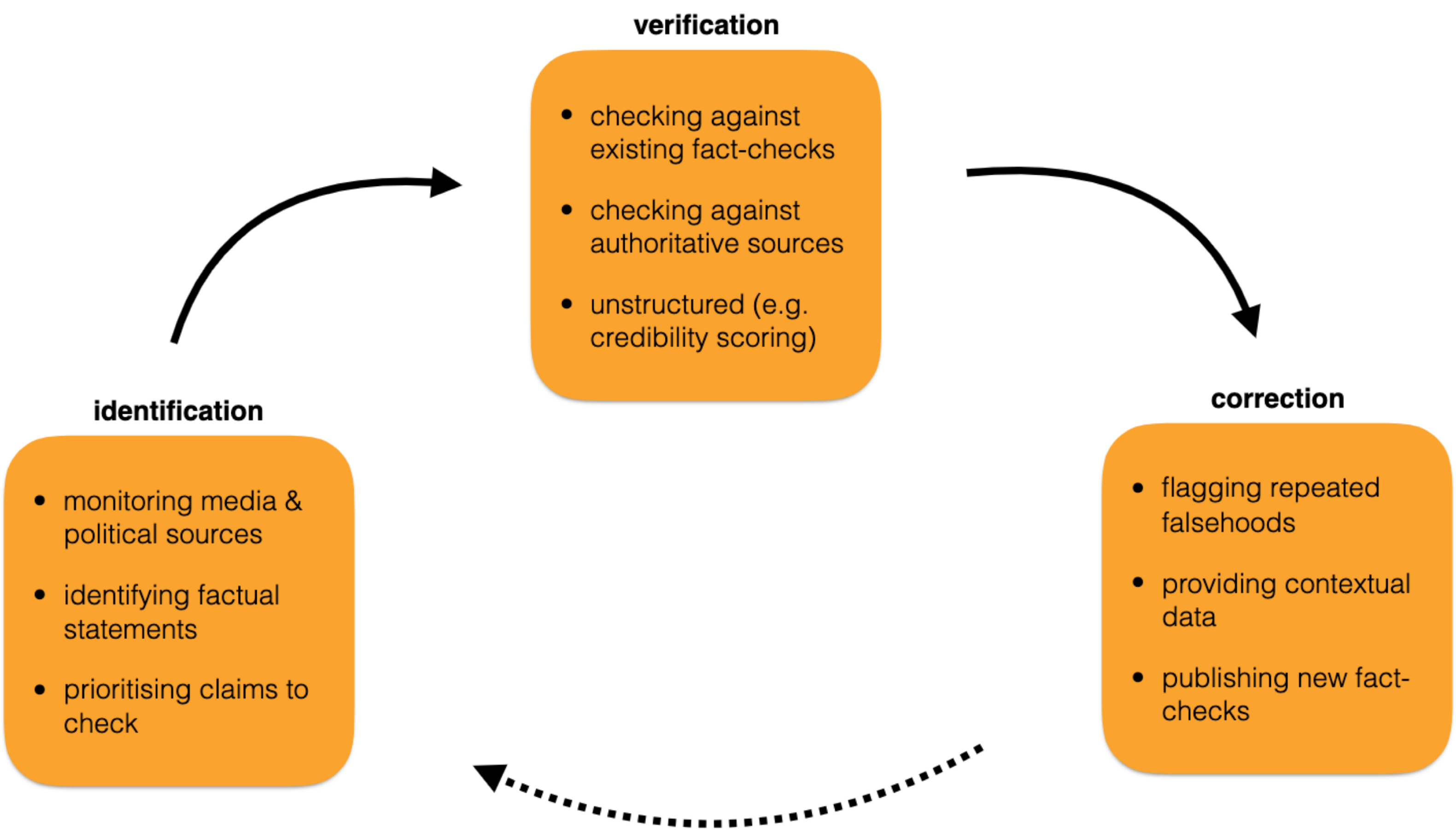
Hackathons, Bootcamps, Labs



AVIGNON **CLEF 2018**

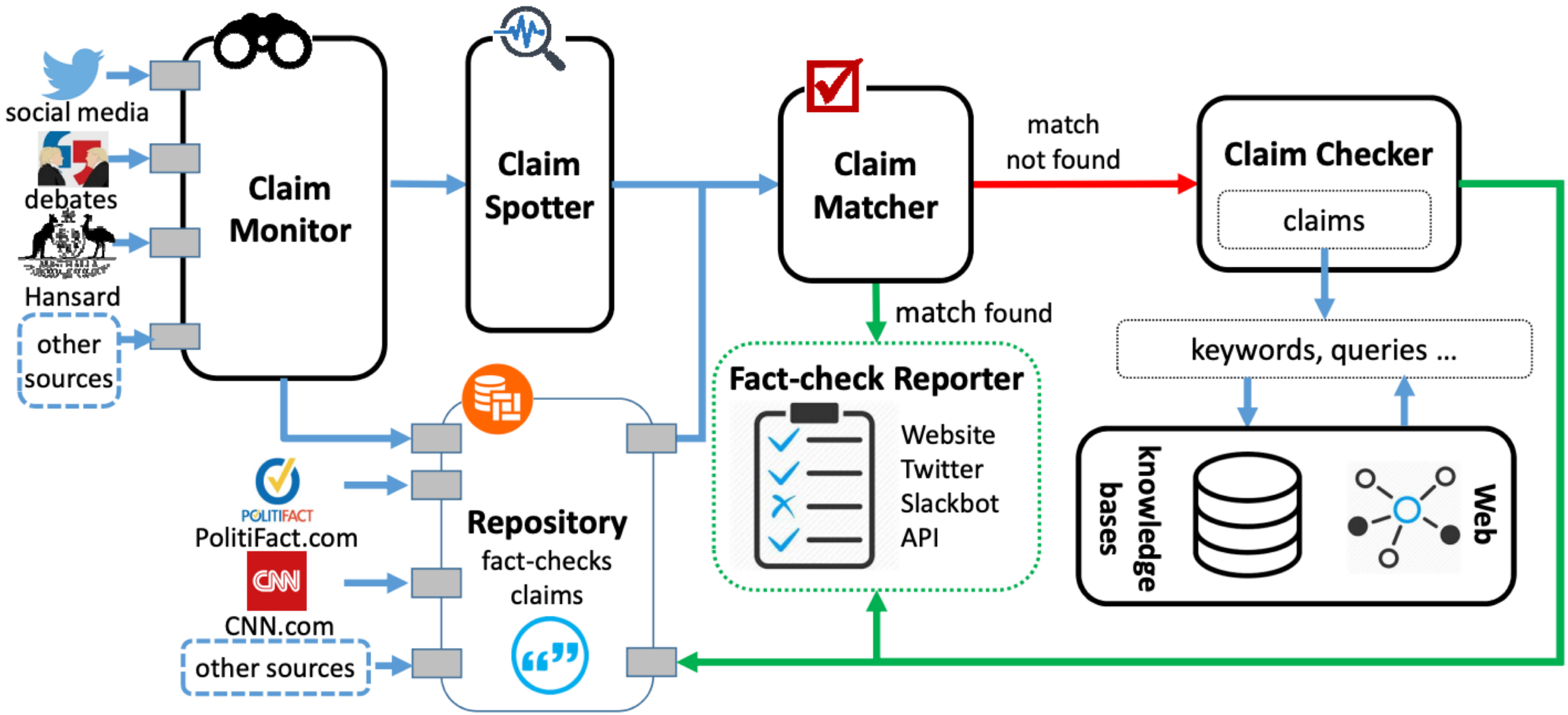
Conference and Labs of the Evaluation Forum

Automated Fact Checking Systems



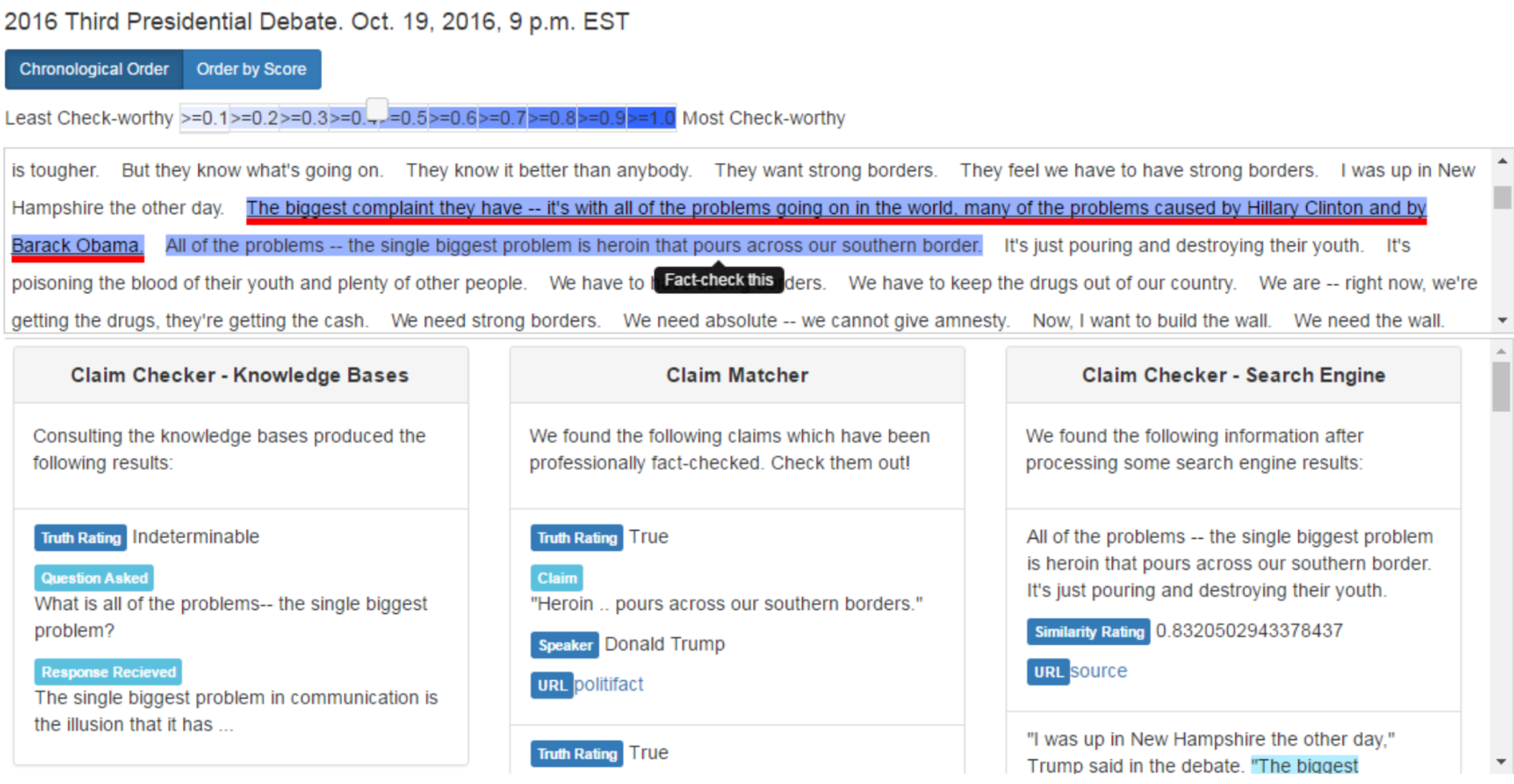
[Lucas Graves. *FactSheet*. 2018]

ClaimBuster Architecture



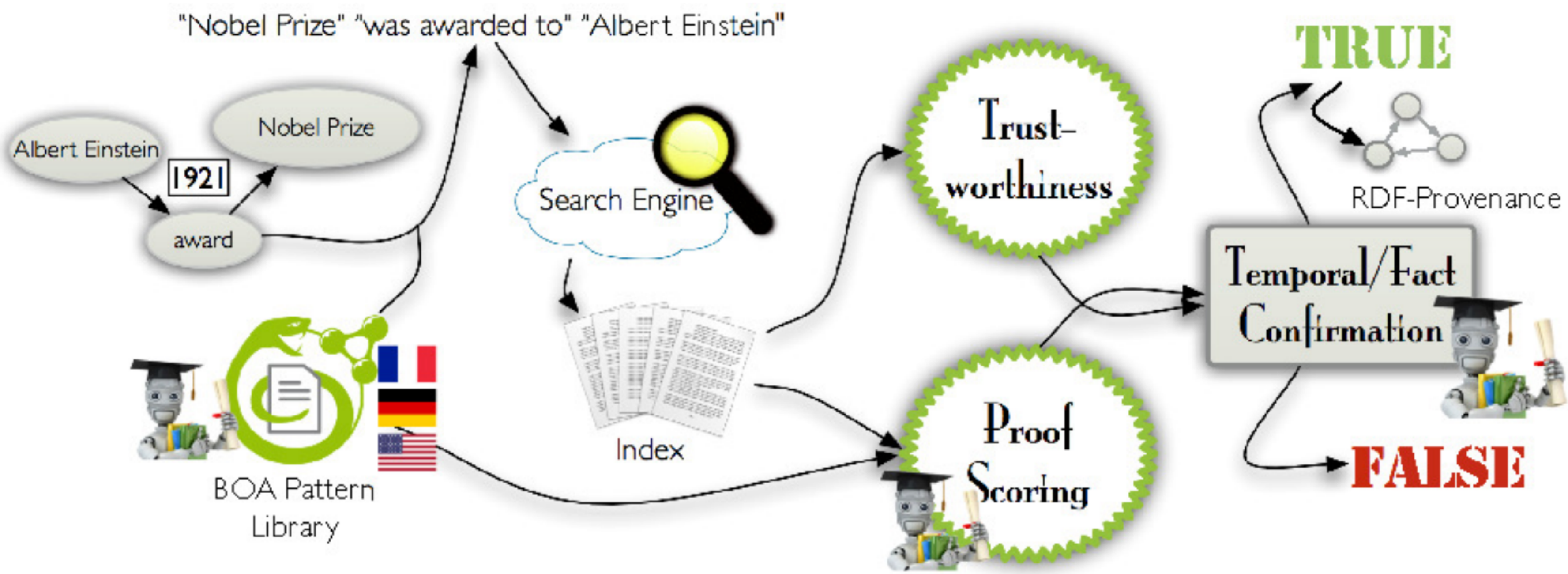
[Hassan et al. *VLDB* 2017.]

ClaimBuster Interface



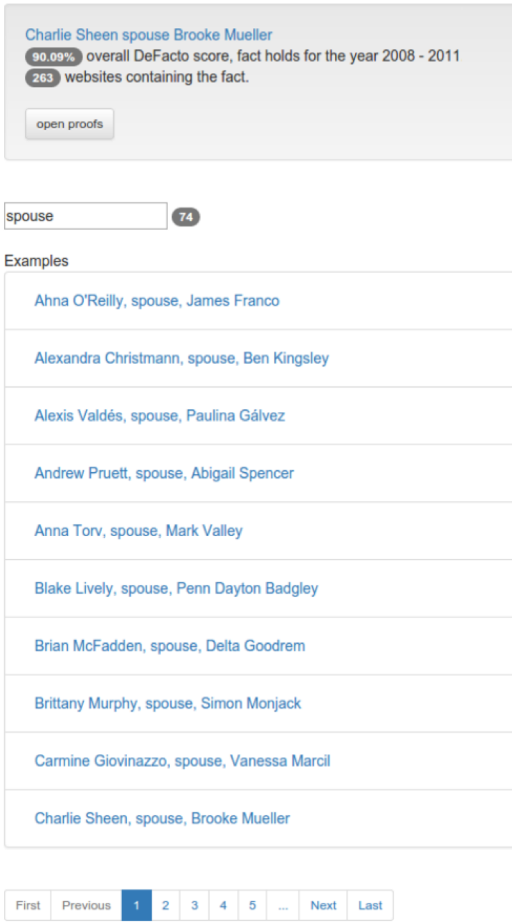
[Hassan et al. *VLDB* 2017.]

DeFacto Functionality

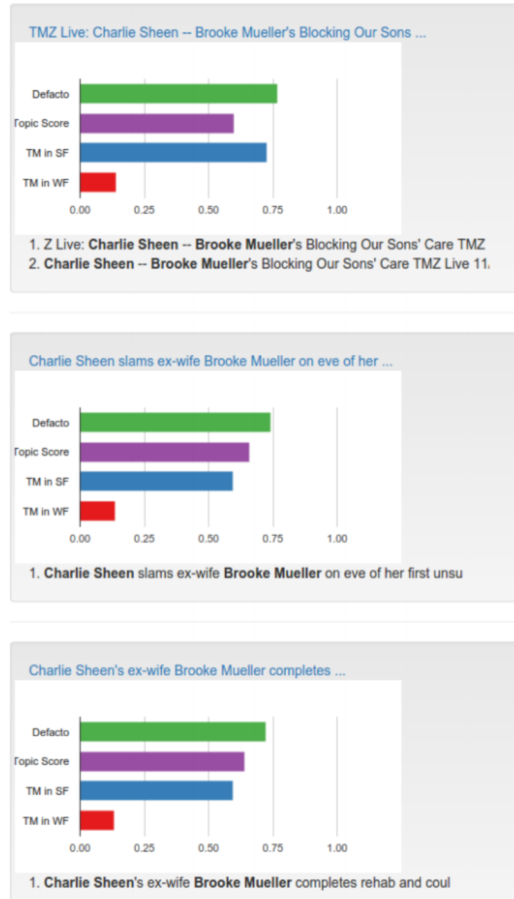


[Speck et al. *ISWC* 2015.]

DeFacto Interface



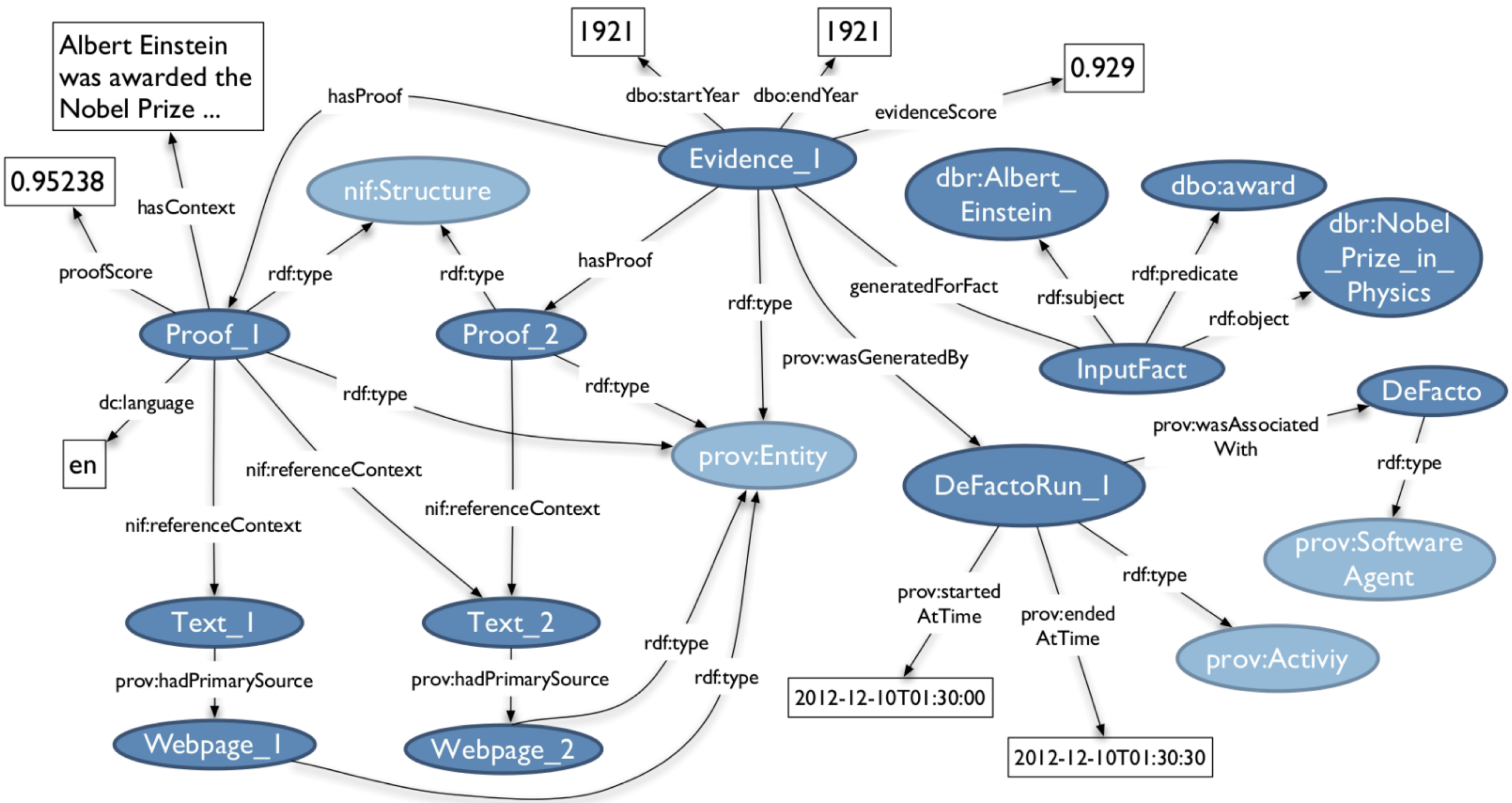
(a) Search form.



(b) Result list.

[Speck et al. *ISWC* 2015.]

DeFacto Evidence and Provenance



[Speck et al. *ISWC* 2015.]

6. Future Challenges & Opportunities

Future Opportunities

- Propagation; Detection; Mitigation; Intervention
- Can DB tech. play a helpful role in Fact Checking?

Key Dimensions

Propagation

Detection

Fake News

Mitigation

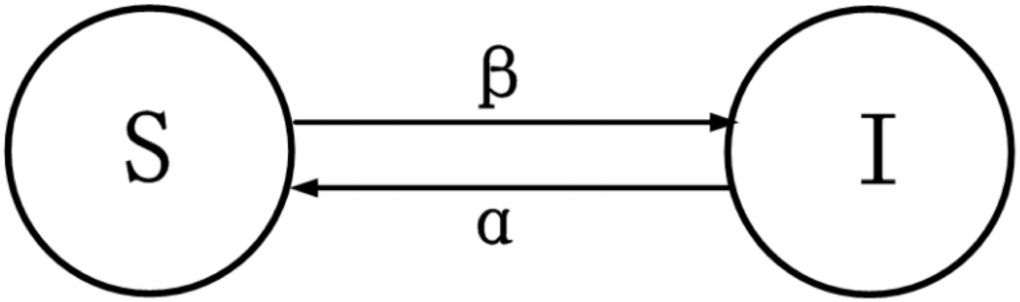
Intervention

Modeling Propagation

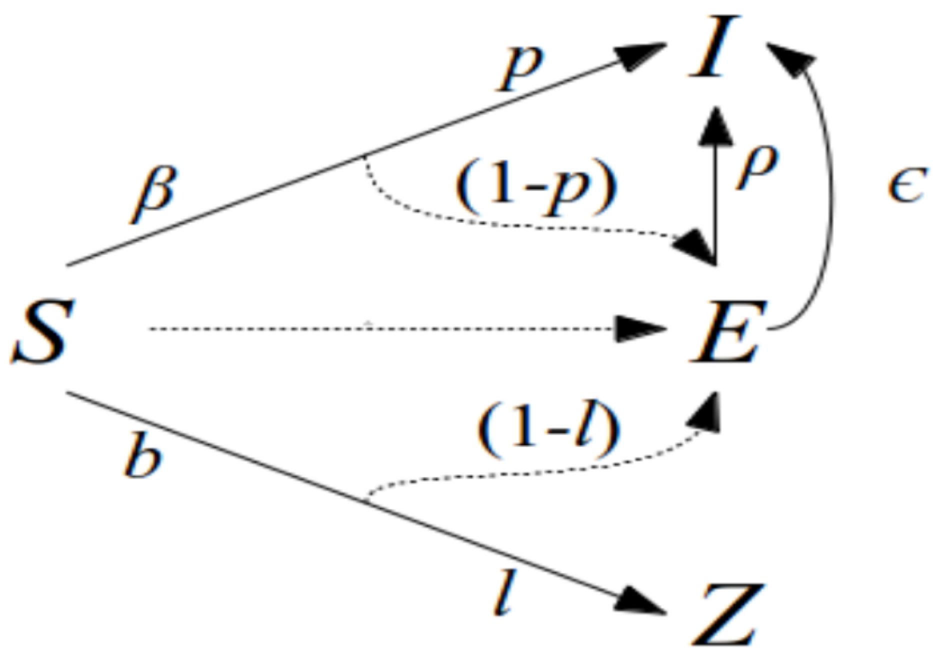
- How to model propagation of fake news?
- SEIZ approach
  - Susceptible: not heard the news
  - Exposed: heard the news and *might* share it
  - Infected: heard the news and already shared it
  - Skeptic: heard the news and did not share

[Fang Jin et al. Epidemiological modeling of news and rumors on twitter. 2013]

Modeling Propagation



SIS Model



SEIZ Model

 [Fang Jin et al. Epidemiological modeling of news and rumors on twitter. 2013].

Empirical Modeling

- Can/should we take a model free approach? How?
- How to empirically model existing fake and real news cascades?

Detection

- Knowledge based Detection
  - Next generation of ML based models have to incorporate knowledge in addition to features
  - How to incorporate KB/KG into a ML classifier?
  - How to integrate ML into query processing based fact checking?

Detection

- Knowledge graph based Fake News Detection
  - Facts as triples stored in KG
  - Popular approach: Link prediction
  - How to generalize from edge to subgraph?

Fact Checking Queries

- Queries are the bread-and-butter of DB community
- How to translate fact checking as queries?
- Are there any novel class of fact checking queries?

Query Perturbation

Claim: “adoptions went up 65 to 70 percent” in New York City “when he was the mayor.”

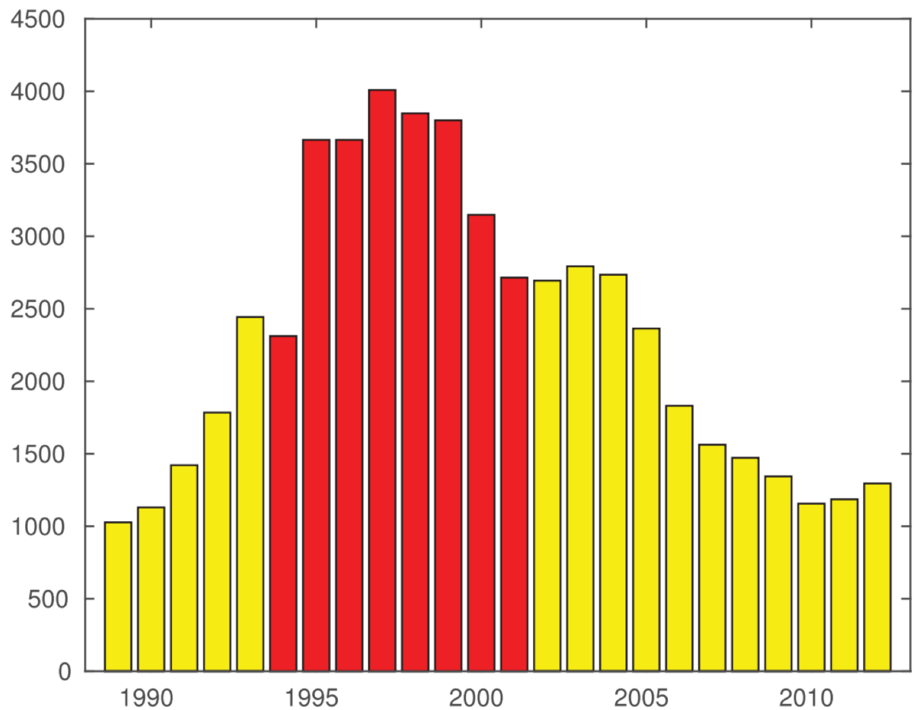


Fig. 1. New York City adoptions by year, 1989–2012. Giuliani’s years in red.

[You Wu et al. "Computational fact checking through query perturbations." TODS. 2017].

Query Perturbation

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;

)

[You Wu et al. "Computational fact checking through query perturbations." TODS. 2017].

Probabilistic Databases

- Real world data is often uncertain and inconsistent
- Can we model fact checking as an inference problem?
  - How to combine uncertain evidence to make decision on fact checking?
- How can we collect and present the evidence for explanations?

[Ahmadi et al. Explainable Fact Checking with Probabilistic Answer Set Programming.].

Mitigation

- How to extend influence maximization and epidemiological models for more accurate mitigation?
  1. Users changing their mind --> switching campaigns
  2. Reacting to evolving propagation --> online setting
- How is seed budget determined in practice?

Intervention

- Should all edges/nodes be treated equally?
  1. Avoid removing highly influential users?
  2. All-or-nothing removal of edges?
- Consider "classes" of users?
  1. E.g. scored w.r.t. their track record (and predicted future credibility)
- How far can the idea of outsourcing intervention via guardians be pushed?
- What other "tagging" schemes are useful?

Fact Checking and DB Technology

Fact Checking Data Warehouse

Ingestion

Fact Checking

Querying Fact Checks

Analytics

Fact Checking Ecosystem

- There is a rich, diverse and thriving ecosystem
- How can our community make the largest impact?
- Build monolithic tools? (use it or lose it?)
- Build specialized tools using "DB" techniques?
- can we redo, what relational did for the enterprise, to fact checking or more generally to truth management?

Questions?

Comments?

References

Introduction

- Allcott and Gentzkow. Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives* 31(2), 2017.
- Wu, Agrawal, Li, Yang, and Yu. Computational Fact-Checking through Query Perturbations. ACM TODS 2017.
- Guo and Vargo, "Fake News" and Emerging Online Media Ecosystem: An Integrated Intermedia Agenda-Setting Analysis of the 2016 U.S. Presidential Election. *Communications Research*, June 2018.

Fake News Propagation

- David Kempe, Jon Kleinberg, and Eva Tardos. Maximizing the spread of influence through a social network. *KDD* 2003.
- Sejeong Kwon, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. Prominent features of rumor propagation in online social media. *ICDM* 2013.
- Soroush Vosoughi, Deb Roy and Sinan Aral. The spread of true and false news online. *Science* 2018.
- Xinyi Zhou, Reza Zafarani. Fake News: A Survey of Research, Detection Methods, and Opportunities. arXiv preprint. 2018.

Data Integration, Truth Discovery & Fusion

- Jing Gao, Qi Li, Bo Zhao, Wei Fan, and Jiawei Han. Truth discovery and crowdsourcing aggregation: A unified perspective. *PVLDB* 2015.
- Yannis Katsis, Yannis Papakonstantinou. View-based data integration. *Encyclopedia of Database Systems*. 2009.
- Theodoros Rekatsinas, Manas Joglekar, Hector Garcia-Molina, Aditya Parameswaran, and Christopher Ré. Slimfast: Guaranteed results for data fusion and source reliability. *SIGMOD* 2017.
- Xin Luna Dong, Evgeniy Gabrilovich, Geremy Heitz, Wilko Horn, Kevin Murphy, Shaohua Sun, and Wei Zhang. From data fusion to knowledge fusion. *PVLDB* 2014.
- Xin Luna Dong, Laure Berti-Equille, and Divesh Srivastava. Integrating conflicting data: the role of source dependence. *PVLDB* 2009.

ML-based Detection

- Subhabrata Mukherjee and Gerhard Weikum. Leveraging Joint Interactions for Credibility Analysis in News Communities. *CIKM* 2015.
- SVN Vishwanathan, Nicol N Schraudolph, Risi Kondor, and Karsten M Borgwardt. Graph kernels. *JMLR* 2010.
- Xinyi Zhou, Reza Zafarani, Kai Shu, and Huan Liu. Fake news: Fundamental theories, detection strategies and challenges. *WSDM* 2019.
- Xinyi Zhou, Reza Zafarani. Fake News: A Survey of Research, Detection Methods, and Opportunities. *arXiv preprint*. 2018.
- Ke Wu, Song Yang, and Kenny Q. Zhu. False rumors detection on sina weibo by propagation structures." *ICDE* 2015.

Knowledge Graph-based Approaches

- Akrami, Farahnaz, et al. "Re-evaluating Embedding-Based Knowledge Graph Completion Methods." *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. ACM, 2018.
- Bordes, Antoine, et al. "Translating embeddings for modeling multi-relational data." *Advances in neural information processing systems*. 2013.
- Chang, Lijun, et al. "Optimal enumeration: Efficient top-k tree matching." *Proceedings of the VLDB Endowment* 8.5 (2015): 533-544.
- Cheng, Jiefeng, Xianggang Zeng, and Jeffrey Xu Yu. "Top-k graph pattern matching over large graphs." *2013 IEEE 29th International Conference on Data Engineering (ICDE)*. IEEE, 2013.
- Ciampaglia, Giovanni Luca, et al. "Computational fact checking from knowledge networks." *PloS one* 10.6 (2015): e0128193.
- Hamilton, Will, et al. "Embedding logical queries on knowledge graphs." *Advances in Neural Information Processing Systems*. 2018.
- Jeh, Glen and Jennifer Widom. SimRank: a measure of structural-context similarity. KDD (2002).

Knowledge Graph-based Approaches

- Kazemi, Seyed Mehran, and David Poole. "Simple embedding for link prediction in knowledge graphs." *Advances in Neural Information Processing Systems*. 2018.
- Lao, Ni, and William W. Cohen. "Relational retrieval using a combination of path-constrained random walks." *Machine learning* 81.1 (2010): 53-67.
- Lin, Peng, et al. "Discovering graph patterns for fact checking in knowledge graphs." *International Conference on Database Systems for Advanced Applications*. Springer, Cham, 2018.
- Lin, Yankai, et al. "Learning entity and relation embeddings for knowledge graph completion." *Twenty-ninth AAAI conference on artificial intelligence*. 2015.
- Lü, Linyuan, Ci-Hang Jin, and Tao Zhou. "Similarity index based on local paths for link prediction of complex networks." *Physical Review E* 80.4 (2009): 046122.
- Morales, Camilo, et al. "MateTee: A semantic similarity metric based on translation embeddings for knowledge graphs." *International Conference on Web Engineering*. Springer, Cham, 2017.

Knowledge Graph-based Approaches

- B. Shi and T. Weninger. Discriminative predicate path mining for fact checking in knowledge graphs. Knowledge-Based Sys., 104:123–133, 2016.
- Shi, Baoxu, and Tim Weninger. "ProjE: Embedding projection for knowledge graph completion." *Thirty-First AAAI Conference on Artificial Intelligence*. 2017.
- P. Shiralkar, A. Flammini, F. Menczer, and G. L. Ciampaglia. Finding streams in knowledge graphs to support fact checking. In 2017 IEEE ICDM 2017, pp 859–864, 2017.
- Wang, Zhen, et al. "Knowledge graph embedding by translating on hyperplanes." *Twenty-Eighth AAAI conference on artificial intelligence*. 2014.
- Xu, Zhongqi, Cunlai Pu, and Jian Yang. "Link prediction based on path entropy." *Physica A: Statistical Mechanics and its Applications* 456 (2016): 294-301.
- Yang, Bishan, et al. "Embedding entities and relations for learning and inference in knowledge bases." *arXiv preprint arXiv:1412.6575* (2014).
- Yang, Shengqi, et al. "Schemaless and structureless graph querying." *Proceedings of the VLDB Endowment* 7.7 (2014): 565-576.
- Yang, Shengqi, et al. "Fast top-k search in knowledge graphs." *2016 IEEE 32nd international conference on data engineering (ICDE)*. IEEE, 2016.

Fake News Mitigation & Intervention

- Bettencourt, Luís MA, et al. "The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models." *Physica A: Statistical Mechanics and its Applications* 364 (2006): 513-536.
- Bharathi, Shishir, David Kempe, and Mahyar Salek. "Competitive influence maximization in social networks." *International workshop on web and internet economics*. Springer, Berlin, Heidelberg, 2007.
- Budak, Ceren, Divyakant Agrawal, and Amr El Abbadi. "Limiting the spread of misinformation in social networks." *Proceedings of the 20th international conference on World wide web*. ACM, 2011.
- Kempe, David, Jon Kleinberg, and Éva Tardos. "Maximizing the spread of influence through a social network." *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2003.
- Khalil, Elias Boutros, Bistra Dilkina, and Le Song. "Scalable diffusion-aware optimization of network topology." *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014.

Fake News Mitigation & Intervention

- Konstantinou, Loukas, Ana Caraban, and Evangelos Karapanos. "Combating Misinformation Through Nudging.” *Co-Inform Project*. 2019.
- Medya, Sourav, Arlei Silva, and Ambuj Singh. "Influence Minimization Under Budget and Matroid Constraints: Extended Version." *arXiv preprint arXiv:1901.02156* (2019).
- Nguyen, Nam P., et al. "Containment of misinformation spread in online social networks." *Proceedings of the 4th Annual ACM Web Science Conference*. ACM, 2012.
- Prakash, B. Aditya, et al. "Threshold conditions for arbitrary cascade models on arbitrary networks." *Knowledge and information systems* 33.3 (2012): 549-575.
- Prakash, B. Aditya, et al. "Fractional immunization in networks." *Proceedings of the 2013 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, 2013.
- Tong, Guangmo, et al. "An efficient randomized algorithm for rumor blocking in online social networks." *IEEE Transactions on Network Science and Engineering* (2017).

Fake News Mitigation & Intervention

- Tong, Hanghang, et al. "On the vulnerability of large graphs." *2010 IEEE International Conference on Data Mining*. IEEE, 2010.
- Vo, Nguyen, and Kyumin Lee. "The rise of guardians: Fact-checking url recommendation to combat fake news." *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 2018.
- Zhang, Yao, and B. Aditya Prakash. "Dava: Distributing vaccines over networks under prior information." *Proceedings of the 2014 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, 2014.
- Zhang, Yao, and B. Aditya Prakash. "Data-aware vaccine allocation over large networks." *ACM Transactions on Knowledge Discovery from Data (TKDD)* 10.2 (2015): 20.
- Zhao, Laijun, et al. "SIHR rumor spreading model in social networks." *Physica A: Statistical Mechanics and its Applications* 391.7 (2012): 2444-2453.

Fact Checking Systems

- Lucas Graves. Understanding the promise and limits of automated fact-checking. *Factsheet* 2018.
- Naeemul Hassan, Gensheng Zhang, Fatma Arslan, Josue Caraballo, Damian Jimenez, Siddhant Gawsane, Shohedul Hasan et al. ClaimBuster: the first-ever end-to-end fact-checking system. *PVLDB* 2017.
- Rene Speck, Diego Esteves, Jens Lehmann, and Axel-Cyrille Ngonga Ngomo. Defacto-a multilingual fact validation interface. *ISWC* 2015.