



CLAIMBUSTER

Toward the Holy Grail of Automated Fact-checking

Chengkai Li

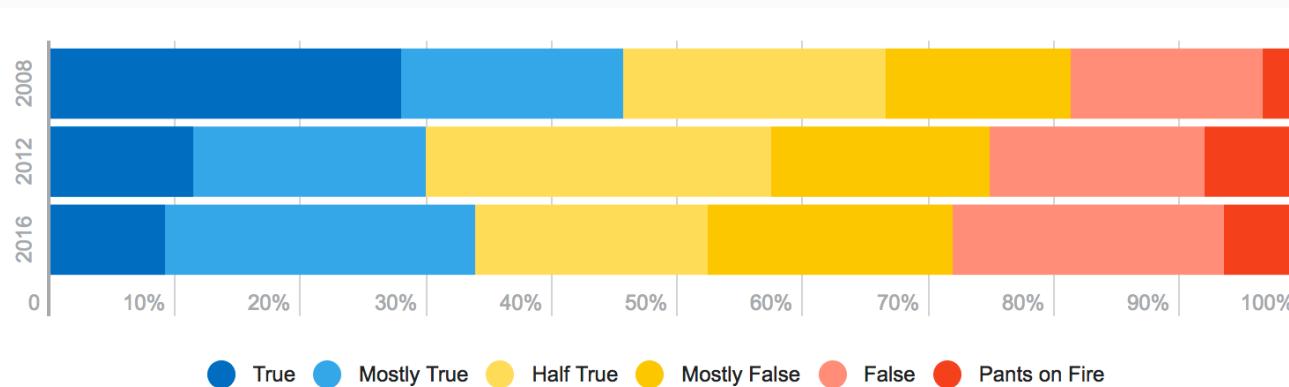
Department of Computer Science and Engineering
University of Texas at Arlington

Google, NYC, May 3rd, 2017



Our Society's Struggle with Falsehoods

- Politicians make false and misleading claims all the time.



As of July 31, 2016

Fact-checks on major party presidential nominees.
Source: PolitiFact.com

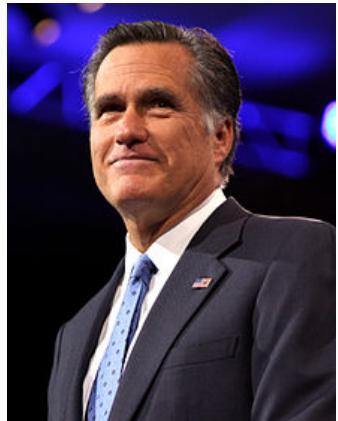


Our Society's Struggle with Falsehoods

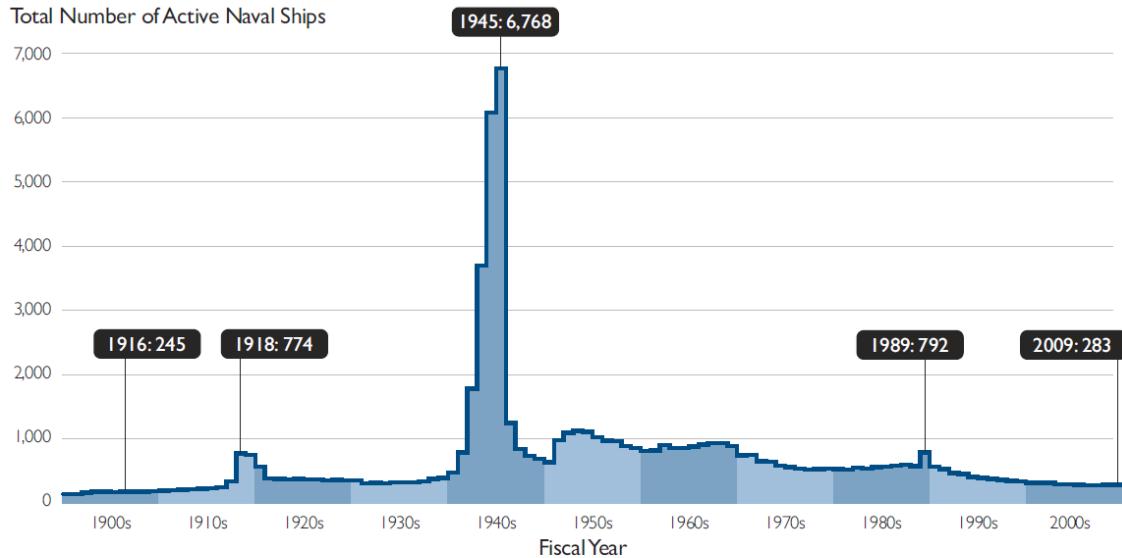
- Politicians make false and misleading claims all the time.
- Fake-news floods the Internet; exacerbated by social media (“filter bubbles” and “echo chambers”).
 - Facebook trend topic algorithms promoted fake-news.
 - A sample of 140,000 Twitter users in the battleground state of Michigan shared as many junk news items as professional news during the final ten days of the 2016 election. <http://politicalbots.org/?p=1064>
- National security threats:
 - U.S. intelligence community concluded that the Russian government interfered with the 2016 election. Fake-news websites and bots used.
 - “Pizzagate”

Fact-checking is Hard

“... our Navy is smaller than it's been since 1917”, said Republican candidate Mitt Romney in third presidential debate in 2012.



U.S. Navy Has Smallest Number of Ships Since 1916

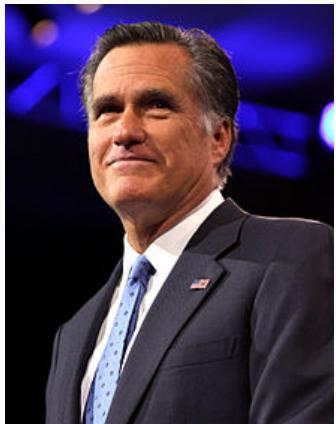


http://en.wikipedia.org/wiki/Mitt_Romney

http://s3.amazonaws.com/thf_media/2010/pdf/Military_chartbook.pdf

Fact-checking is Hard

“... our Navy is smaller than it's been since 1917”, said Republican candidate Mitt Romney in third presidential debate in 2012.



VS



http://en.wikipedia.org/wiki/Mitt_Romney

http://s3.amazonaws.com/thf_media/2010/pdf/Military_chartbook.pdf

http://en.wikipedia.org/wiki/United_States_Navy

Fact-Checking is Time-Consuming and Resource-Demanding



≡ THE NEW YORKER ☐


BOROWITZ REPORT
JANUARY 16, 2016

FACT CHECKER AT REPUBLICAN DEBATE HOSPITALIZED FOR EXHAUSTION

BY ANDY BOROWITZ





Satire



Combating Falsehoods

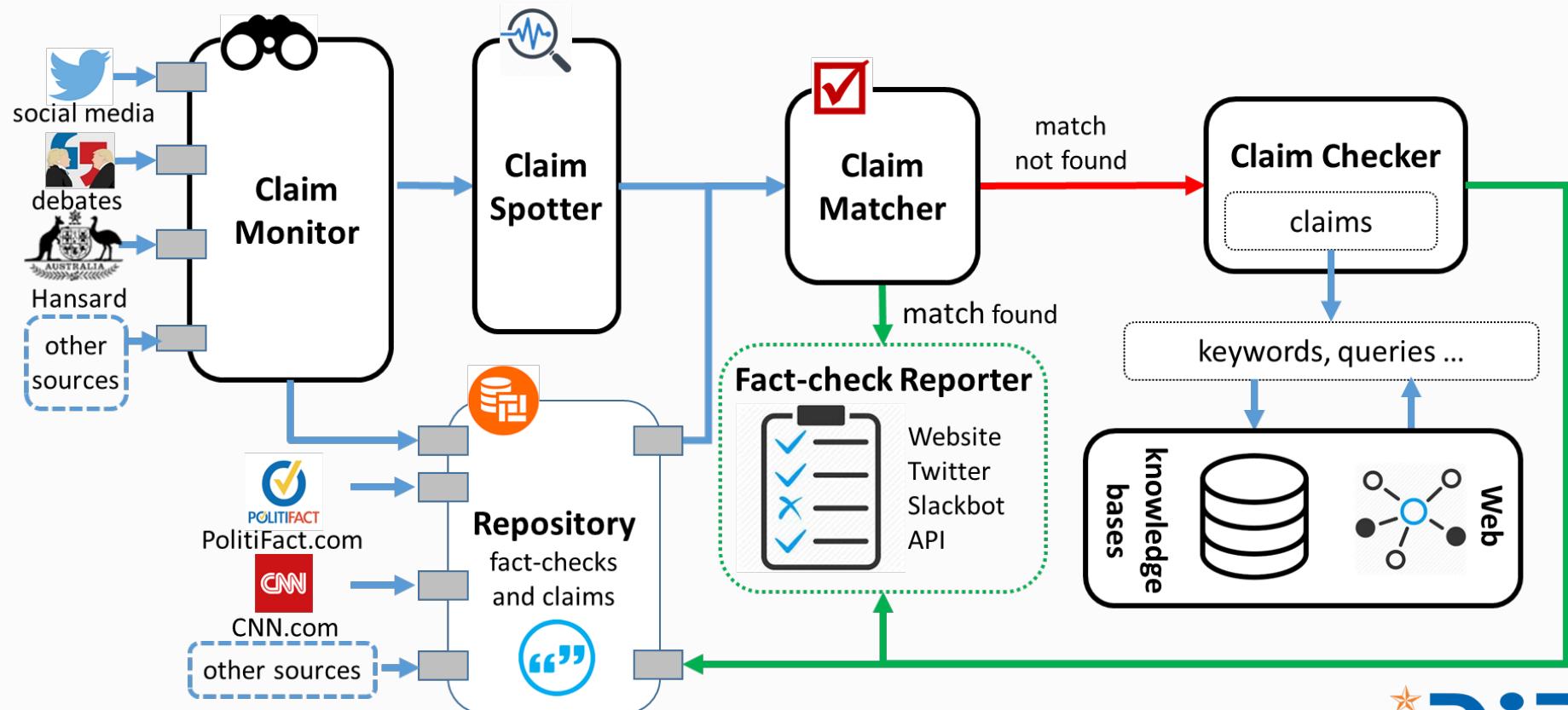
- Initiatives by Google, Facebook, and others
 - 100+ active fact-checking sites in 2017 (PolitiFact.com, FullFact.org, CNN, Washington Post, New York Times, ...)
 - Google includes fact-checks in search results.
 - Facebook lets users report false items and flags items disputed by fact-checkers.
- Three categories of algorithm-based approaches
 - Leveraging experts' verdicts
 - Trait-based (rumor detection)
 - Data-driven, algorithmic



CLAIMBUSTER



The First-ever End-to-end System

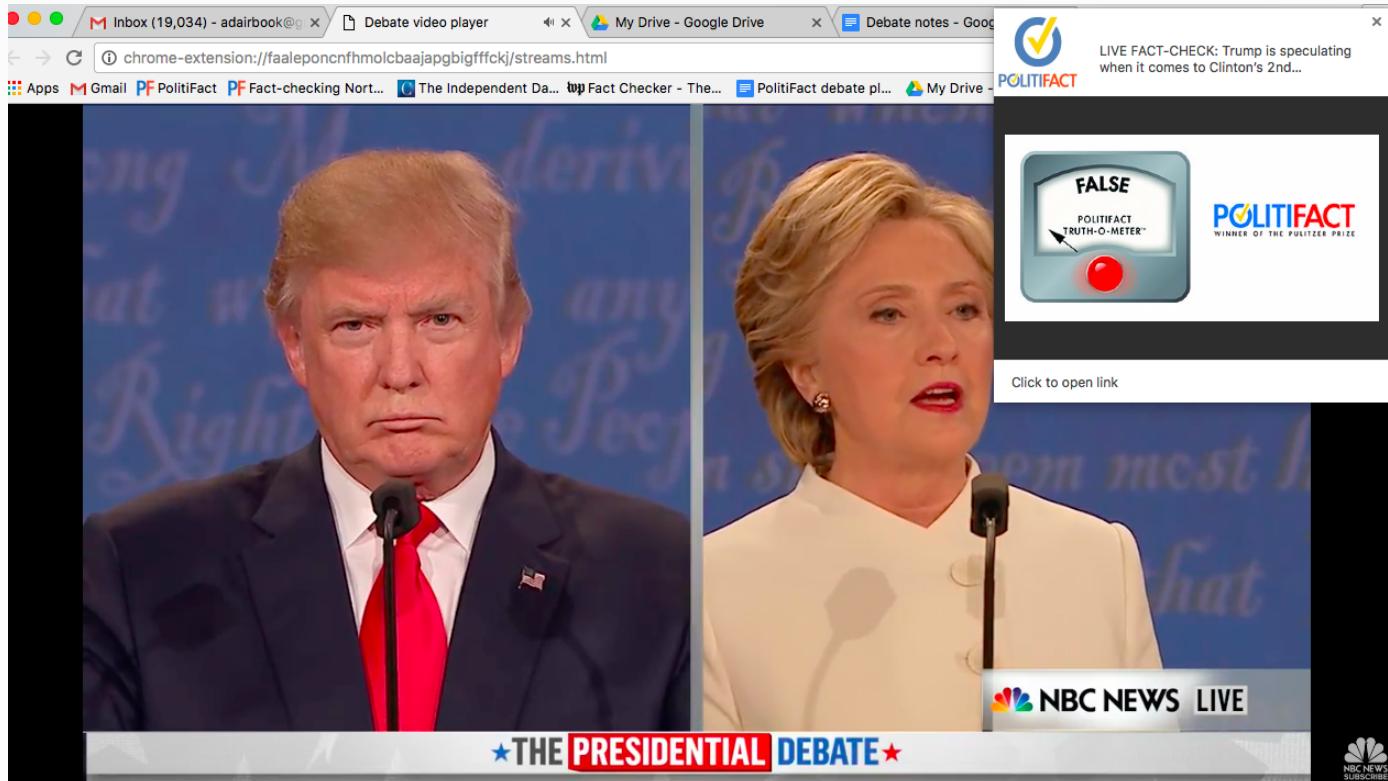




The Holy Grail: Automated, Live Fact-Checking



Proof of Concept: PolitiFact Fact-check Popup



Source: Duke Reporters' Lab



Enter Your Own Text



2016 U.S. Presidential Debates



Hansard: Parliament of Australia



ClaimBuster Slackbot

End-to-end Fact-checking (beta)

Tweets by @ClaimBusterTM

ClaimBuster Retweeted



Rep. Susan Davis

@RepSusanDavis



San Diego gets about \$800 million in @NIH funding each year. Proud to lead the effort to get a \$2 billion increase in this year's budget.



6h

ClaimBuster Retweeted



Governor Walker

@GovWalker



Cumulative tax relief - including tax cuts in our **#WIbudget** - over 8 years exceeds \$8 billion! Tell lawmakers to lower your taxes!



Embed

View on Twitter

**ON AIR****2016 Third Presidential Debate Live. Oct. 19, 2016, 9 p.m. EST**

Participants: Donald Trump, Hillary Clinton

Moderators: Chris Wallace

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

Participants: Hillary Clinton, Donald Trump

Moderators: Chris Wallace

**2016 Second Presidential Debate. Oct. 9, 2016, 9 p.m. EST**

Participants: Hillary Clinton, Donald Trump

Moderators: Martha Raddatz, Anderson Cooper, Chris Wallace

**2016 Vice Presidential Debate. Oct. 4, 2016, 9 p.m. EST**

Participants: Tim Kaine, Mike Pence, Hillary Clinton, Donald Trump

Moderators: Elaine Quijano, Martha Raddatz, Anderson Cooper

**2016 First Presidential Debate. Sep. 26, 2016, 9 p.m. EST**

Participants: Hillary Clinton, Donald Trump

Moderators: Lester Holt

**2016 Democratic Party Presidential Debate. April 14, 2016, 9 p.m. EST**

Participants: Bernard Sanders, Hillary Clinton, Donald Trump

CLAIM TO CHECK

Drinking champagne could help prevent Alzheimer's and other forms of dementia.

**Claim Checker - Knowledge Bases**

Consulting the knowledge base produced the following results:

INCLUDE THIS IN REPORT

Question: What could help prevent Alzheimer's and other forms of dementia?

Google Answer Box: People also ask

Details**Generate a report from selected evidence****Claim Matcher**

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

INCLUDE THIS IN REPORT **RATING** False**CLAIM**

Drinking champagne could help prevent Alzheimer's and other forms of dementia.

URL [politifact](#)**INCLUDE THIS IN REPORT** **RATING** false**CLAIM**

Red Bull and other brands of energy drinks contain bull semen.

URL [snopes](#)**Claim Checker - Search Engine**

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

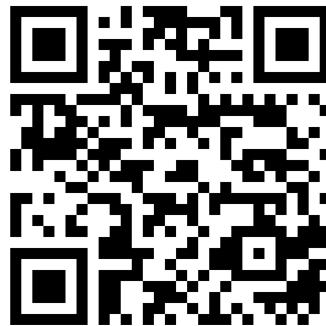
INCLUDE THIS IN REPORT

Nov 9, 2015 - Scientists have found that drinking three glasses of champagne every day can help to prevent the onset of dementia and Alzheimer's.

SIMILARITY RATING 0.556**URL** [cosmopolitan](#)**INCLUDE THIS IN REPORT**

Sep 1, 2016 - Drinking champagne may help prevent dementia and Alzheimer's ... other things, the neurotransmitters essential for memory recall can function ...

SIMILARITY RATING 0.554**URL** [nowtolove.com](#)



ClaimBuster APP 5:08 PM Only visible to you

Working on your request.

@gensheng

When my husband was president, we went from a \$300 billion deficit to a \$200 billion surplus and we were actually on the path to eliminating the national debt.

Score

0.657743014341

Working on your request.

@gensheng

We know how to get control of the debt.

Score

0.11061813488983



ClaimBuster + Slack

* ClaimBuster provides a platform for you to evaluate pieces of text and identify sentences that are check-worthy. This is our first step in our long-term goal of automating fact-checking. Our Slack app allows you to score individual sentences using the `/getscore claim` command, create a Dropbox folder on our servers using the `/createfolder e-mail` command, store text files from the web into your assigned Dropbox folder for future processing on our servers using the `/processfile e-mail file-url` command, and submit feedback to help us keep improving our services using the `/claimbusterfeedback message` command.



Add to Slack



ClaimBuster API Documentation

apidocs : ClaimBuster API

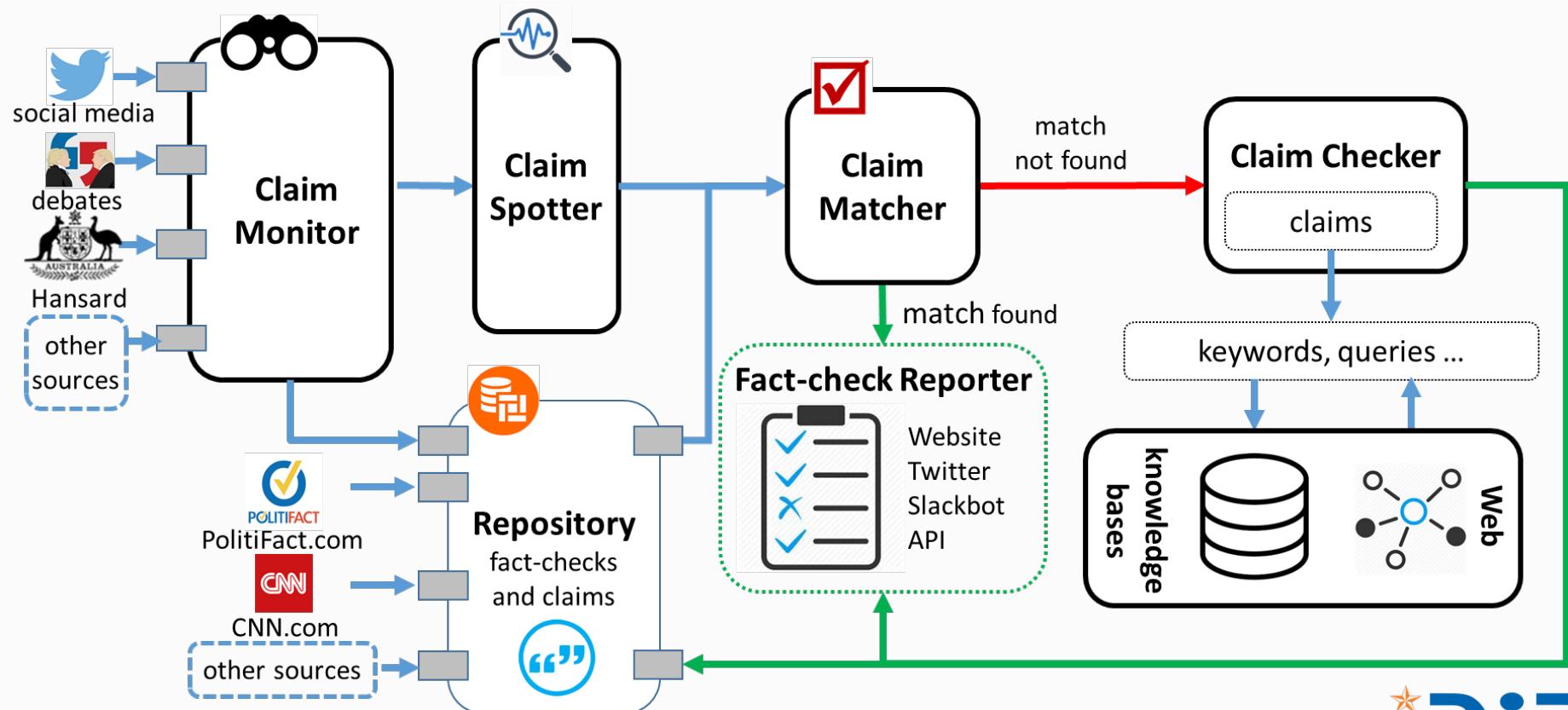
Show/Hide | List Operations | Expand Operations | Raw

GET	/all/{claim}	All
GET	/se/{claim}	Claim Checker - Search Engine
GET	/fm/{claim}	Claim Matcher
GET	/kb/{claim}	Claim Checker - Knowledge Bases
GET	/score_text/{text}	Claim Spotter - Get Score
GET	/score_url/{url}	Claim Spotter - Get Score URL



Claim Spotter

The First-ever End-to-end System



PolitiFact “Buffet” of Factual Claims



	A	B	C	D	E
1	Statement	Speaker	Date	Location/Source	Link
232	I haven't heard one (republican) say that they're in concentrated wealth and capital have been put into it is not true that regulation holds poor people down or	o'malley o'malley o'malley	4/20/2015 morning edition 4/20/2015 morning edition 4/20/2015 morning edition		http://www.npr.org/blogs/its http://www.npr.org/blogs/its http://www.npr.org/blogs/its
235	Every three weeks we bring online as much as solar	Obama	4/18/2015 Weekly address		https://www.whitehouse.gov
236	Our carbon polution has fallen by 10 percent since	Obama	4/19/2015 Weekly address		https://www.whitehouse.gov
237	*Benghazi has had more hearings, more documents	Claire McCaskill	4/19/2015 ABC this week		http://thehill.com/policy/inte
238	With fast track trade " I would be receiving the same	Obama	4/17/2015 remarks joint press		https://www.whitehouse.gov
239	Loretta Lynch "has been now sitting there longer than	Obama	4/17/2015 remarks joint press		https://www.whitehouse.gov
240	I just counted 19 Republicans who are likely to run for	chris cillizza	4/20/2015 the twitters		https://twitter.com/TheFix/st
241	He [Rand Paul] wanted to cut all our defense	John McCain	4/20/2015 Fox News		https://www.youtube.com/w

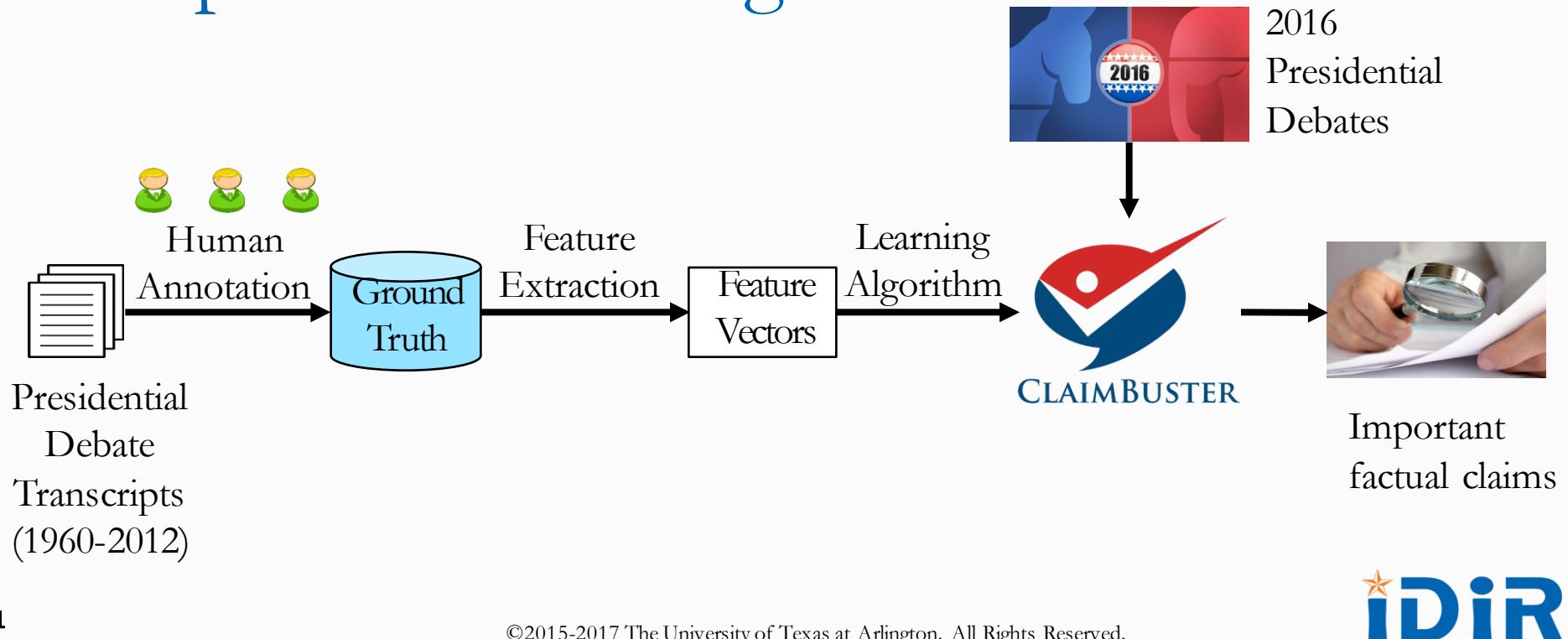
2014

This is what they eat at PolitiFact.



Let me work on its automation.

Finding Important Factual Claims: A Supervised Learning Task



Dataset: Presidential Debate Transcripts

- Source: <http://www.debates.org/index.php?page=debate-transcripts>
- All 30 debates (11 general elections) in history: 1960, 1976—2012
- 20k sentences by presidential candidates: removed very short (< 5 words) sentences



3 Classes of Sentences

CFS: Important factual claims

“We spend less on the military today than at any time in our history.” “The President’s position on gay marriage has changed.” “More people are unemployed today than four years ago.”

UFS: Unimportant factual claims

“I was in Iowa yesterday.” “My mother enjoys cooking.” “I ran for President once before.”

NFS: No factual claims (just opinions, questions & declarations)

“Iran must not get nuclear weapons.” “7% unemployment is too high.” “My opponent is wishy-washy.” “I will be tough on crime.” “Why should we do that?” “Hello, New Hampshire!” “Our plan is to reduce tax rate by 10%.”



3 Classes of Sentences

If a sentence contains an opinion and an important factual claim, it shouldn't be labelled “No factual claims”.

“Unemployment rate is too high and in last four years it increased from 8% to 10%.”

Ranking and Classification Problems



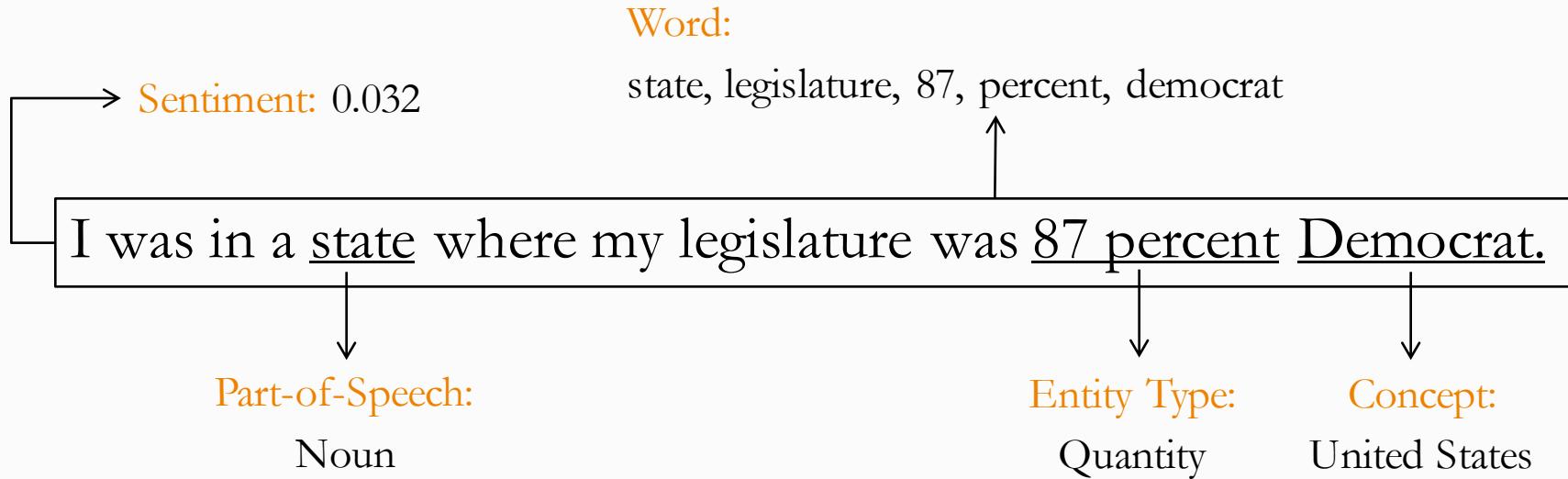
Classification

Given a future sentence, find the class it belongs to.

Ranking

Rank sentences based on how likely they contain important factual claims.

Feature Extraction



Sentiment: [-1.0 to 1.0]

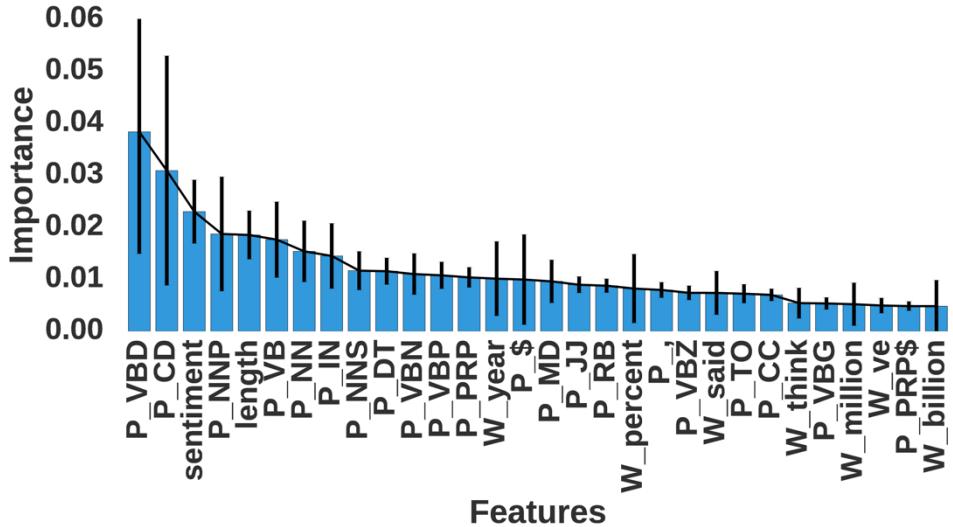
Word: tf-idf scores of 6130 words (excluding rare words)

POS Tag: 43 tags

Entity Type: 26 types

Feature Importance

- 6615 features in total
- Used a Random Forest Classifier to calculate the importance of each feature.



Word	percent; people; jobs
POS Tag	noun; cardinal number; past tense; preposition
Entity Type	Quantity; Country; FieldTerminology; Person
Concept	United States Senate; Barack Obama
Sentiment	



Evaluation: Classification

- 4-fold cross validation
- Methods: Naive Bayes (NBC), Random Forest (RFC), Support Vector Machine (SVM)
- In general, SVM, using Word + POS, performed better than other combinations.

	Precision	Recall	F-measure
NFS	0.90	0.96	0.93
UFS	0.65	0.26	0.37
CFS	0.79	0.74	0.77

Comparison with Subjectivity Classifiers



Table 4: Comparison with [22]

	NFS	UFS	CFS
subjective	157	5	44
objective	574	58	194

Table 5: Comparison with [21]

	NFS	UFS	CFS
subjective	21	0	4
unknown	175	5	45
objective	535	58	189

Evaluation: Ranking

- NBC, RFC, SVM all allow ranking.
- Accuracy of top-K sentences ranked by ClaimBuster, using the admitted sentences from top-coders as ground truth.
- ClaimBuster has a strong agreement with top-quality human coders on the check-worthiness of sentences.

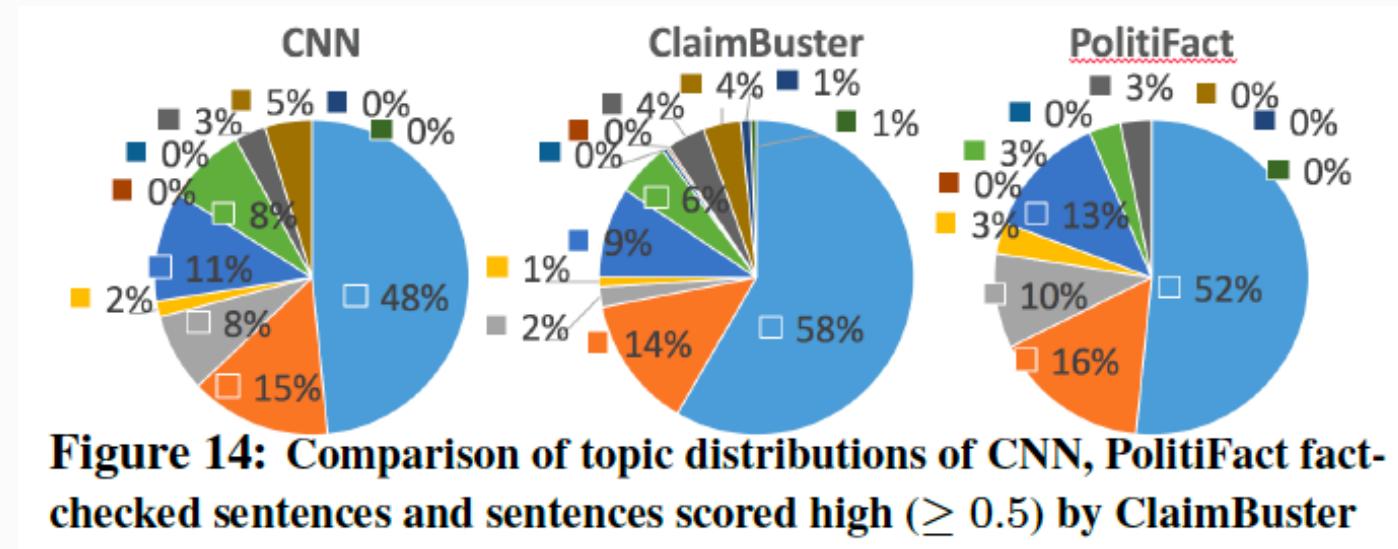
k	P@k	AvgP	nDCG
10	1.000	1.000	1.000
25	1.000	1.000	1.000
50	0.980	0.995	0.987
100	0.943	0.979	0.956
200	0.916	0.955	0.931
300	0.848	0.937	0.874
400	0.764	0.915	0.801
500	0.679	0.897	0.827

Case Study: 2016 U.S. Presidential Election Debates



Table 6: Score differences between sentences fact-checked and those not chosen for checking

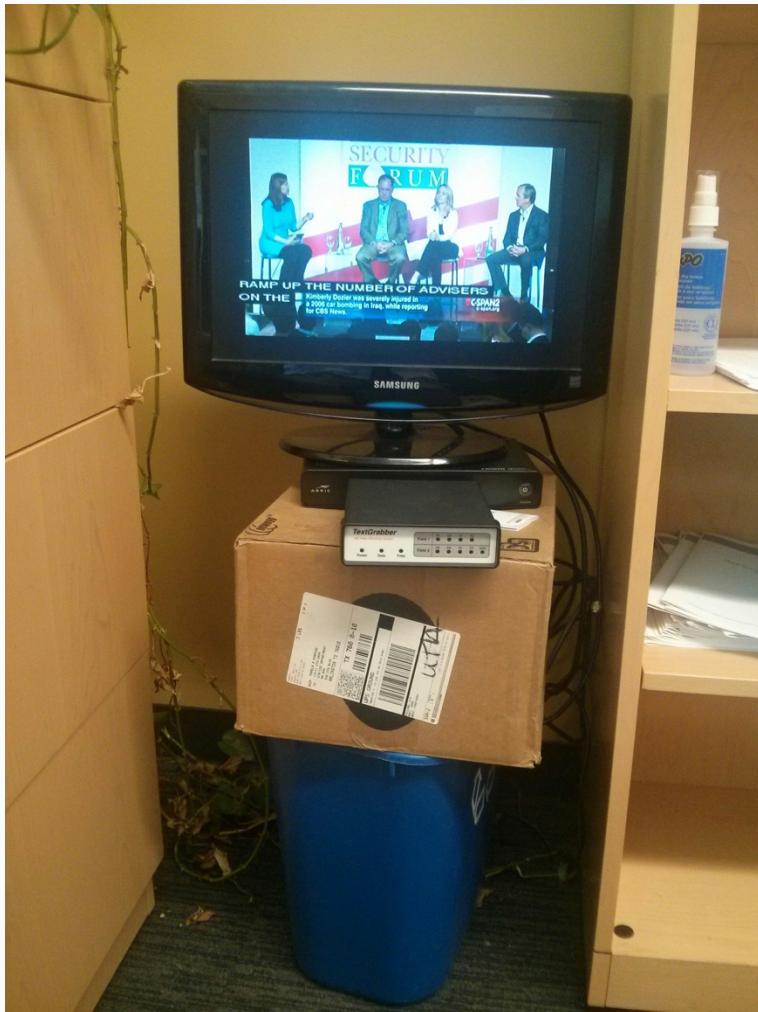
Platforms	avg(YES)	avg(NO)	t-value	p-value
CNN	0.433	0.258	21.137	1.815E-098
PolitiFact	0.438	0.258	16.362	6.303E-060





Live Coverage of Debates

- Extracted closed captions of debates on broadcast TV channels using TextGrabber
- Finding speaker identity using Twitter
- Tweeted important factual claims to
<https://twitter.com/ClaimBusterTM>, live!





Claims Tweeted

Tweets [Follow](#)

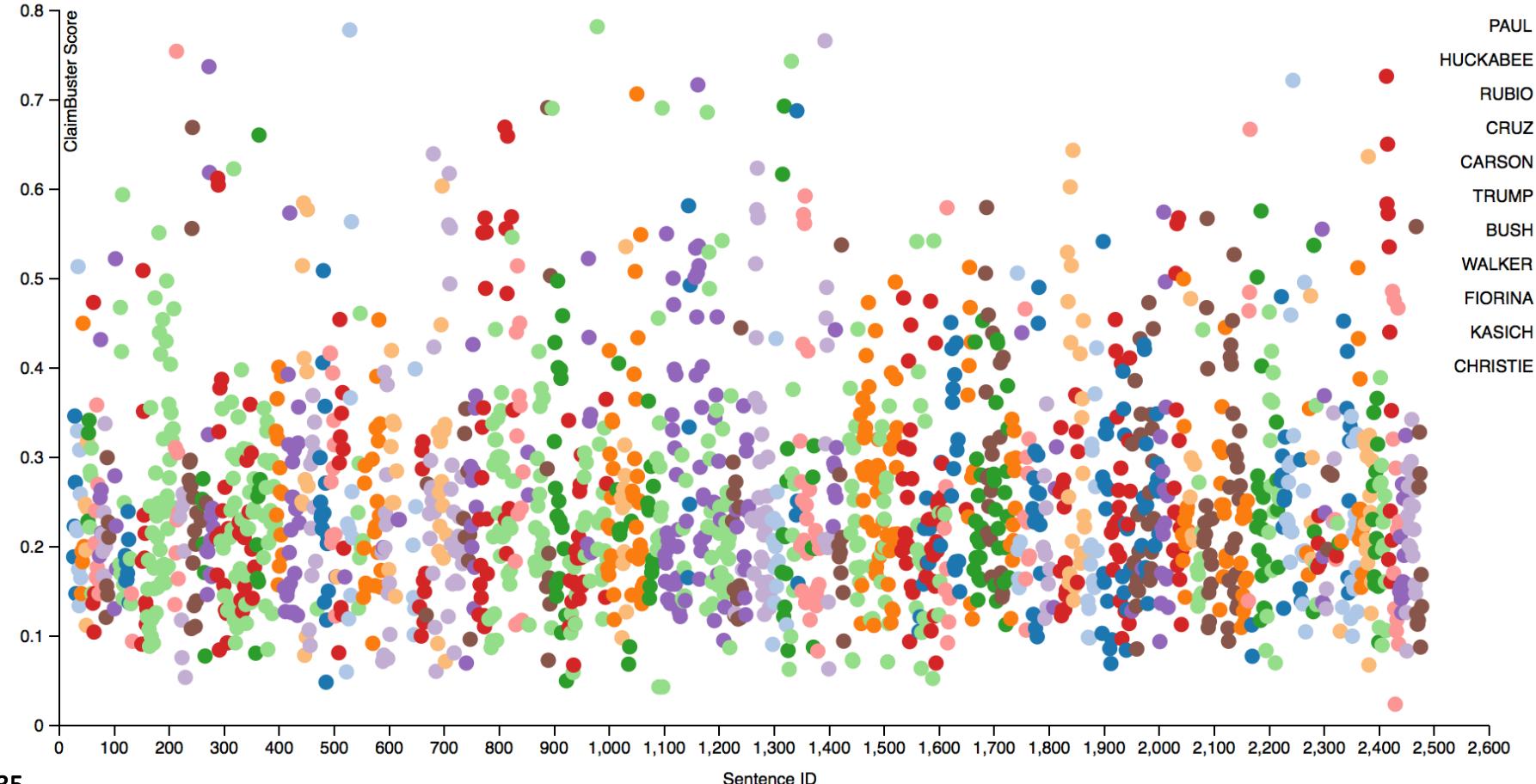
 **ClaimBuster** 14 Nov
@ClaimBusterTM
(1/2) #PolitiFactThis [BERNIE SANDERS?] We have the highest rate of childhood poverty, and we're the only country in the world --

 **ClaimBuster** 14 Nov
@ClaimBusterTM
(2/2) #PolitiFactThis virtually the only country that doesn't guarantee paid family and medical leave.

 **ClaimBuster** 14 Nov
@ClaimBusterTM
(1/2) #PolitiFactThis It wasn't the bill that I wanted, but yet it turned out to be one of the more significant pieces of



Debate Timeline Graph





Retweeting Claims in Tweets

Tweets [Follow](#)

 **ModerateEdge** 3m
@ModerateEdge

@realDonaldTrump If you make \$25,001,
should you pay \$2,500 when \$25,000 you
pay nothing? Pay only on amt over \$25K.
#Trump2016 #DonaldTrump
rt Retweeted by ClaimBuster
[Expand](#)

 **BicycleBrandsDirect** 9m
@bicyclebrands

1.3 million bicycles recalled for crash
hazard: (KMS) - Nearly 1.3 million
bicycles in the United States are ...
bit.ly/1MGT2HO
rt Retweeted by ClaimBuster
[Expand](#)

 **Rich Luchette** 19m
@richluchette

\$272 million project, more than 3,000 jobs
[Tweet to @ClaimBusterTM](#)



Ground-Truth Collection



Our Own CrowdFlower



ClaimBuster

naeemulhassan labeled 562 sentences

[Leaderboard](#)

[Instructions](#)

[Log Out](#)

Your estimated total payment is \$37.39. You are paid approximately 6.70¢ per sentence on average. Your total payment is ranked 8 out of 54 participants. Click 'Leaderboard' to see details. This message and the leaderboard are updated every 15-30 sentences. [See tips for improving your pay rate at the bottom of this page.](#)

RO: Now, I voted for the Patriot Act.

[More Context](#)

Will the general public be interested in knowing whether (part of) this sentence is true or false?

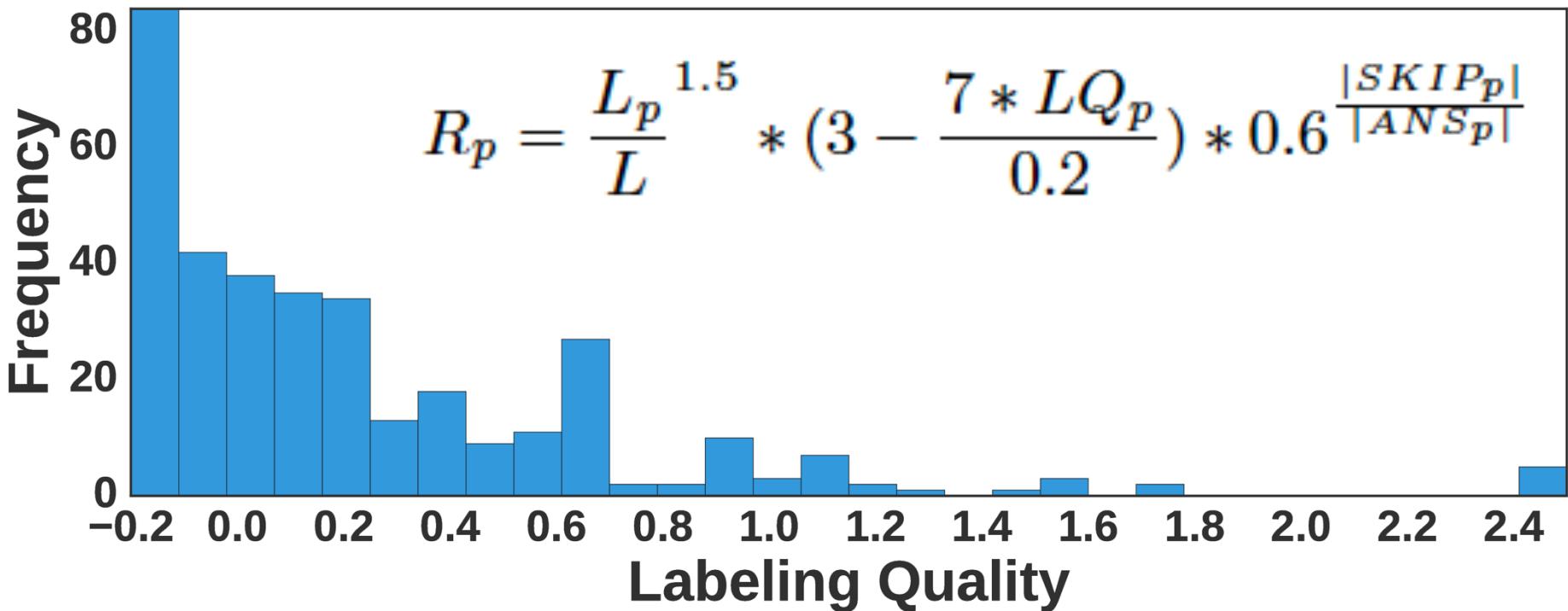
- There is **no** factual claim in this sentence.
- There is a factual claim but it is **unimportant**
- There is an **important** factual claim.

[Submit](#)

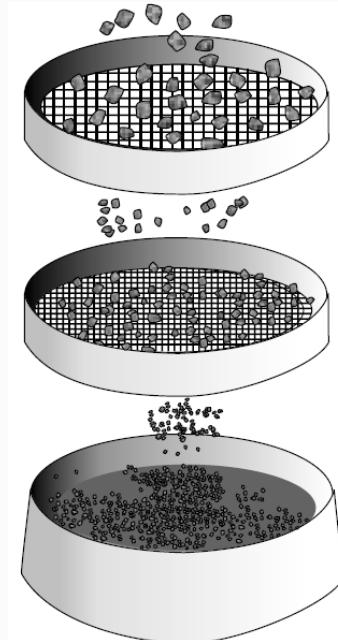
[Skip this sentence](#)

[Modify My Previous Responses](#)

Coder Quality



Ground Truth Collection



376 coders

76552 labels

88 top-quality coders

52333 labels

88 top-quality coders

17665 admitted sentences

- 20 months
- ~\$4,000 paid
- 1032 screening sentences (by 3 experts) to detect spammers & low-quality coders

Class	Count	Percentage
CFS	4187	23.70%
UFS	1627	9.21%
NFS	11851	67.08%

You are Invited to Contribute



<http://bit.ly/claimbusters>

 ClaimBuster

naeemulhassan labeled 562 sentences [Leaderboard](#) [Instructions](#) [Log Out](#)

Your estimated total payment is \$37.39. You are paid approximately 6.70¢ per sentence on average. Your total payment is ranked 8 out of 54 participants. Click 'Leaderboard' to see details. This message and the leaderboard are updated every 15-30 sentences. [See tips for improving your pay rate at the bottom of this page.](#)

RO: Now, I voted for the Patriot Act.

[More Context](#)

Will the general public be interested in knowing whether (part of) this sentence is true or false?

There is **no** factual claim in this sentence.

There is a factual claim but it is **unimportant**

There is an **important** factual claim.

[Submit](#) [Skip this sentence](#) [Modify My Previous Responses](#)

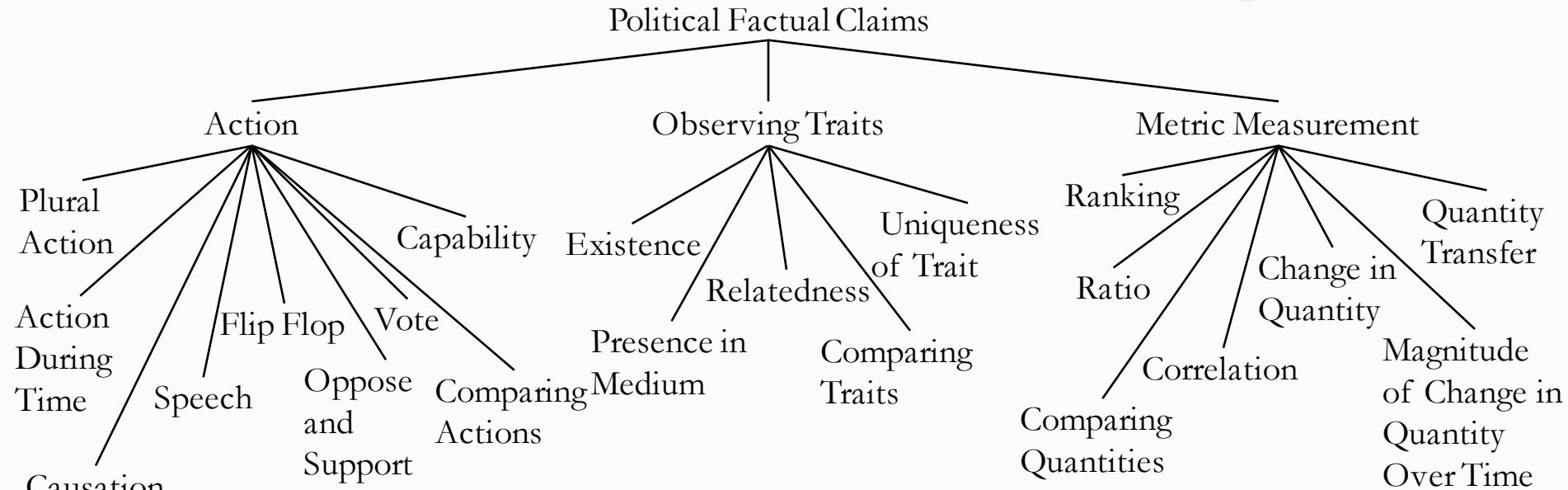




PolitiTax: A Taxonomy of Political Factual Claims

PolitiTax

- Modeling Claims
- Can be useful to all relevant ClaimBuster components





Example: Ranking

E: The entity which is being ranked

GROUP: The group of entities which *E* is being ranked against

VAR: The metric which all entities in *GROUP* have

R: The ranking of *E* against all other entities in *GROUP*

Examples

The United States is $R^{65^{\text{th}}}$ out of $GROUP^{142}$ nations and other territories on $VAR^{equal\ pay}$.

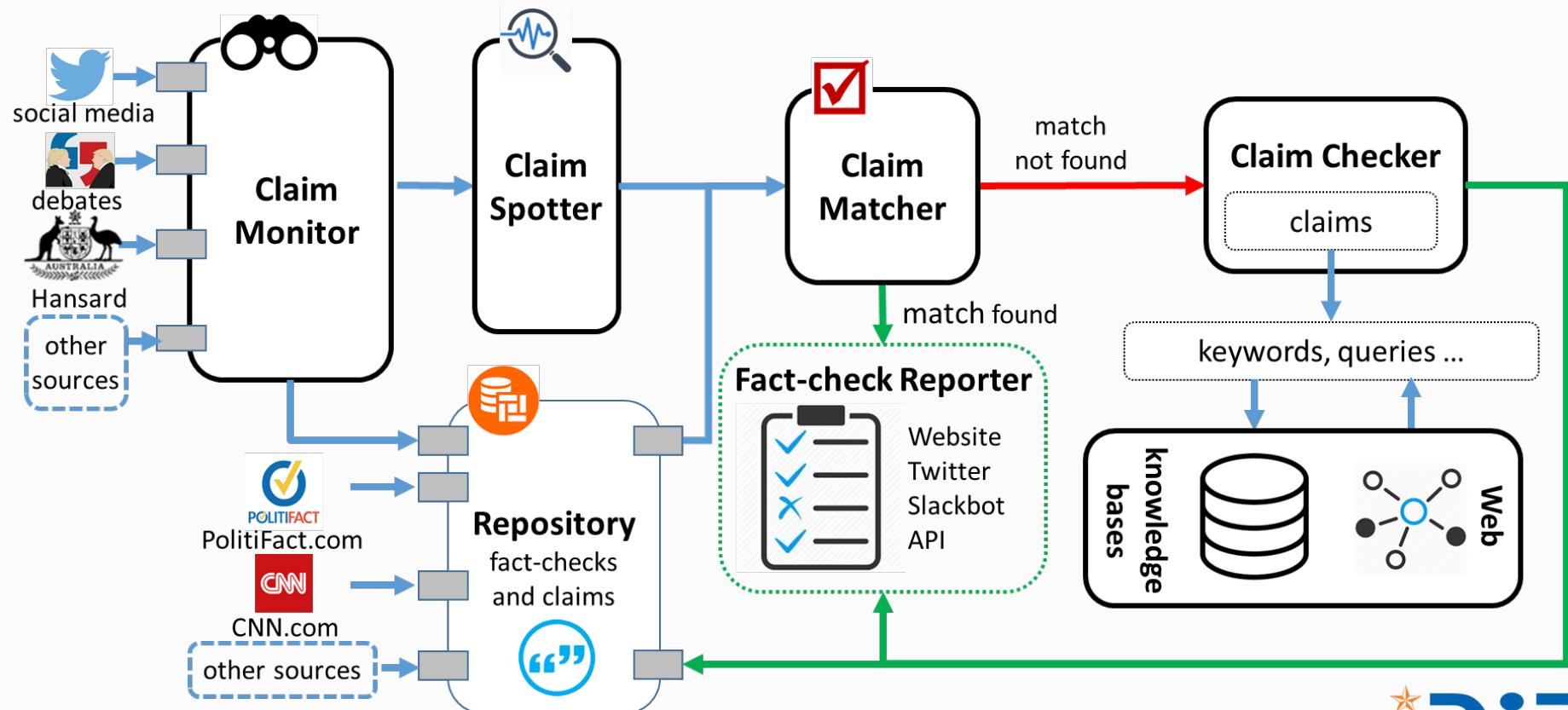
We're now the $R^{No.\ 1}$ $VAR^{producer\ of\ oil}$ in $GROUP^{the\ world}$. We've surpassed Saudi Arabia



Claim Matcher



The First-ever End-to-end System





Textual Entailment Algorithms

Textual entailment is defined as a **directional** relationship between pairs of text expressions, denoted by T (the entailing “Text”) and H (the entailed “Hypothesis”). We say that T entails H if humans reading T would typically infer that H is most likely true.



Preliminary Experiments

- Small dataset:
 - 2500 pairs(text&hypothesis): Cross product
 - Text (50 sentences from 2016 presidential debate)
 - Hypothesis (50 sentences from Politifact)
- Medium dataset:
 - ~7 million pairs(text&hypothesis)
 - Text (all 2016 presidential debate sentences 32k)
 - Hypothesis (only debate related sentences from Politifact ~220)
- Large dataset:
 - ~35 million pairs (text&hypothesis)
 - Text (Politifact sentences ~6k)
 - Hypothesis (2016 presidential debate sentences ~6k)

Preliminary Experiments

- EditDistance

	Precision	Recall
Small Dataset	0.87	1
Medium Dataset	1	0.83

- TIE (MaxEntClassificationEDA)

	Precision	Recall
Small Dataset	1	0.88
Large Dataset	1	Ongoing



Acknowledgment

Current/Former UTA Students

- Naeemul Hassan
- Gensheng Zhang
- Fatma Arslan
- Josue Caraballo
- Damian Jimenez
- Siddhantsanjeev Gawsane
- Shohedul Hasan
- Minumol Joseph
- Aaditya Kulkarni
- Anil Kumar Nayak
- Vikas Sable

Collaborators

- Bill Adair (Duke)
- Pankaj Agarwal (Duke)
- Peter Frays (University of Technology Sydney)
- James Hamilton (Stanford)
- Mark Tremayne (UTA)
- You Wu (Google)
- Jun Yang (Duke)
- Cong Yu (Google)



Acknowledgment

Funding sponsors



UNIVERSITY OF
TEXAS
ARLINGTON

Disclaimer: This material is based upon work partially supported by the National Science Foundation Grants 1408928 and 1565699 and a Knight Prototype Fund from the Knight Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.



Thank You! Questions?

- <http://ranger.uta.edu/~cli>
cli@uta.edu
- Demo

<http://idir.uta.edu/claimbuster>

- Please help us label the data
<http://bit.ly/claimbusters>

