

The Quest to Automate Fact-Checking

Chengkai Li

Associate Professor, Department of Computer Science and Engineering

Director, Innovative Database and Information Systems Research (IDIR) Laboratory

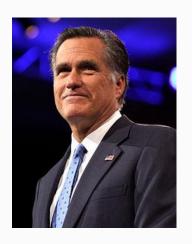
University of Texas at Arlington

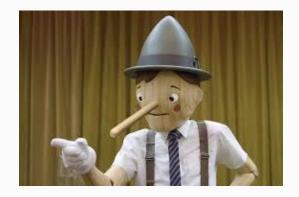






"... our Navy is smaller than it's been since 1917", said Republican candidate Mitt Romney in third presidential debate in 2012.







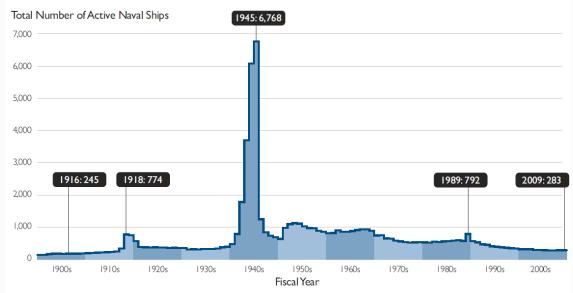




"... our Navy is smaller than it's been since 1917", said Republican candidate Mitt Romney in third presidential debate in 2012.



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



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

Source: U.S. Navy, Active Ship Force Levels, 2009, at http://www.history.navy.mil/branches/org9-4.htm (December 6, 2009).

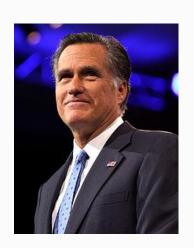
//s3.amazonaws.com/thf_media/2010/pdf/Military_chartbook.pdf

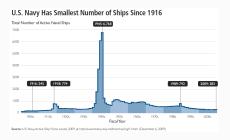


Fact-checking is not Easy



"... 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/United_States_Navy



Existing Fact-checking Platforms



Journalists and reporters spend good amount of time on fact checking



The U.S. military is at risk of losing its "military superiority" because "our Navy is smaller than it's been since 1917. Our Air Force is smaller and older than any time since 1947."



— Mitt Romney on Monday, January 16th, 2012 in a Republican presidential debate in Myrtle Beach, S.C.

96 active fact-checking sites in 2016 (PolitiFact .com, FactCheckEU.org, FullFact.or, Factcheck .org CNN, Washington Post, New York Times, Poynter, American Press Institute, . . .)

https://reporterslab.org/global-fact-checking-up-50-percent/



"Total Physical and Mental Collapse"









BOROWITZ REPORT

JANUARY 16, 2016

FACT CHECKER AT REPUBLICAN DEBATE HOSPITALIZED FOR EXHAUSTION

BY ANDY BOROWITZ







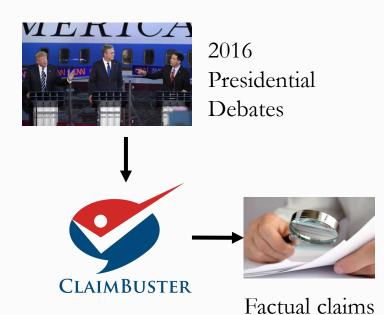












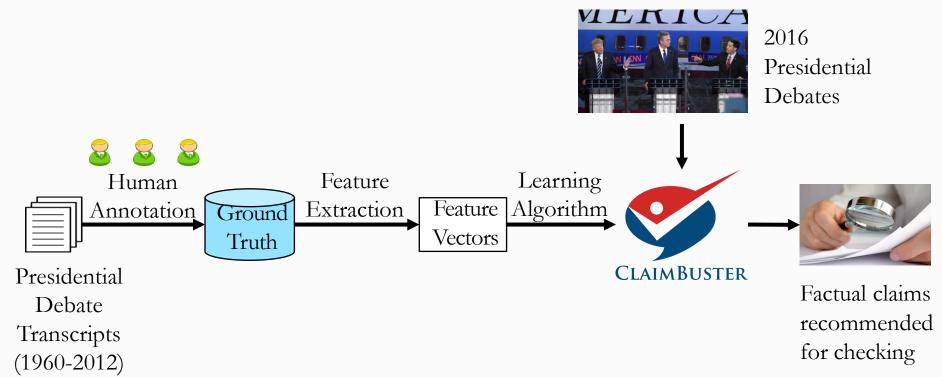


recommended

for checking



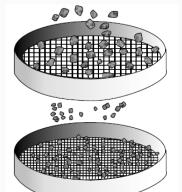
Fact-Checking Presidential Debates





Ground Truth Collection



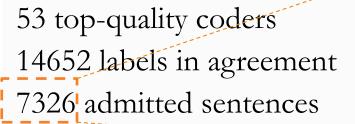


329 coders

51925 labels

53 top-quality coders 26853 labels

- o 10 months (in 3 stages)
- o ~\$3,000 paid
- 1032 screening sentences (by
 3 experts) to detect spammers
 & low-quality coders



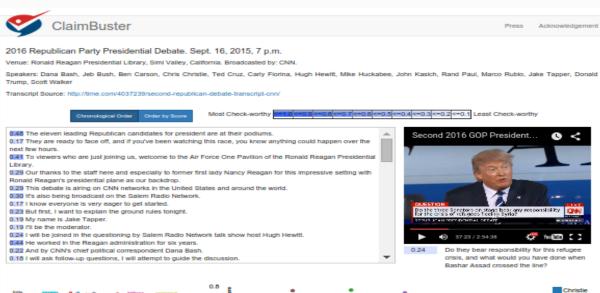
Class	Count
CFS	1561
UFS	418
NFS	5347



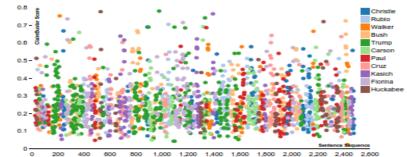
Live Coverage of Debates













ClaimBuster automated live fact-checking

Fact-check your own text



2016 Democratic Party Presidential Debate. Jan. 17, 2016, 9 p.m. ET

Paticipants: Hillary Clinton, Bernard Sanders, Martin O'Malley

Moderators: Lester Holt, Andrea Mitchell



2016 Republican Party Presidential Debate. Jan. 14, 2016, 9 p.m. ET

Paticipants: Ted Cruz, John Kasich, Chris Christie, Jeb Bush, Marco Rubio, Ben Carson, Donald Trump

Moderators: Neil Cavuto, Maria Bartiromo



2016 Democratic Party Presidential Debate. Dec. 19, 2015, 8 p.m. ET

Paticipants: Hillary Clinton, Martin O'Malley, Bernard Sanders Moderators: David Muir, Martha Raddatz, Josh McElveen



2016 Republican Party Presidential Debate. Dec. 15, 2015, 9 p.m. ET

Paticipants: Rand Paul, John Kasich, Chris Christie, Carly Fiorina, Jeb Bush, Marco Rubio, Ted Cruz, Ben

Carson, Donald Trump

Moderators: Wolf Blitzer, Hugh Hewitt, Dana Bash

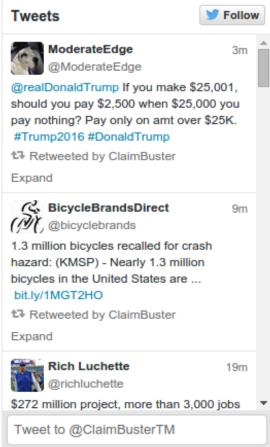


2016 Democratic Party Presidential Debate. Nov. 14, 2015, 8 p.m. CST



Retweeting Claims in Tweets







The Holy Grail: Automated, Live Fact-Checking





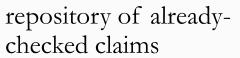




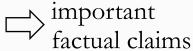
ClaimBuster: What to be Done

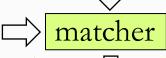
- live events (debates, interviews, congressional debates) and ads on TV and online video streams
- social media (e.g., twitter)
- web pages
- news articles

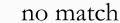






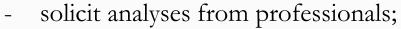






matched





- crowd;
- data analysis, algorithmic factchecking



display existing fact-checkers, delivered via browser extensions, mobile and smart-TV apps



Media Coverage















FIRST DRAFT

Austin American-Statesman







Commercialization







Application Areas

- o Journalism
- o Public relations
- o Opposition research
- Vetting Crowdfounding projects







80K (GRA, faculty salary, data collection, materials)

for 2 years

to develop some of the components

for one of the applications



Thank You! Questions?

- o http://ranger.uta.edu/~cli cli@uta.edu
- Prototypehttp://idir.uta.edu/claimbuster
- You are invited to contribute http://bit.ly/claimbusters



