



## C+J2015 Submission 19


[Update authors](#)

If you want to **change any information** about your paper or withdraw it, use links in the upper right corner.

For all questions related to processing your submission you should contact the conference organizers. [Click here to see information about this conference.](#)

All **reviews sent to you** can be found at the bottom of this page.

### Paper 19

Title:	The Quest to Automate Fact-Checking
Paper	
Author keywords:	Computational Journalism Fact-checking Presidential Election Debate Machine Learning Natural Language Processing
EasyChair keyphrases:	fact checking (300), fact check (140), factual claim (110), fact checker (80), holy grail (70), presidential debate (70), data source (60), natural language (50), check worthy claim (47), support vector machine (47), time consuming (40), numeric value (40), computational journalism (40), closed caption (40)
Abstract:	Our purpose with this paper is twofold: To argue for the advancement of research on automated fact-checking and to report our progress toward that end with a tool called ClaimBuster developed via machine learning. We describe the current state of fact checking research and describe the approach we've taken with ClaimBuster. We report preliminary results of a field test comparing ClaimBuster's ability to identify check-worthy factual claims made during the August, 2015, Republican presidential debate with those of professional journalists and fact-checking organizations. Significant overlaps between the machine scoring and human judgments were observed.
Time:	Aug 14, 23:56 GMT

### Authors

first name	last name	email	country	organization	Web site	corresponding?
Naeemul	Hassan	naeemul.hassan@mavs.uta.edu	United States of America	University of Texas at Arlington		✓
Bill	Adair	bill.adair@duke.edu	United States of America	Duke University		✓
James	Hamilton	jayth@stanford.edu	United States of America	Stanford University		✓
Chengkai	Li	cli@uta.edu	United States of America	University of Texas at Arlington		✓
Mark	Tremayne	tremayne@uta.edu	United States of America	University of Texas at Arlington		✓
Jun	Yang	junyang@cs.duke.edu	United States of America	Duke University		✓
Cong	Yu	congyu@google.com	United States of America	Google Research		✓

## Reviews

### Review 2

The paper deals with political fact-checking. When is a politicians claim a fact? There are so many language problems here -- The first is, what constitutes a claim? Their tool, ClaimBuster, is an attempt to help journalists find claims to fact-check. They provide a useful classification of statements into Non-Factual Sentences, Unimportant Factual Sentences, and Check-worthy Factual Sentences. Their Review: classification system is meant to spot the CFSS. They then tested their system on the 2015 GOP Debate.

I like this paper a lot. There are so many problems I hadn't considered were at all approachable. It's a first start and I think that makes it interesting to talk about in our forum.

### Review 1

This paper addresses the need for automated checking of political facts. I like this paper. This isn't a computational task that I had thought about previously, but it's clear that there is a clear need for such tools. Although I believe that the authors' "holy grail" for automated fact-checking is unachievable in the near future (given state-of-the-art in image/video/text analysis and in machine learning), I really liked their clear articulation of this ideal. I also liked their detailed list of computational challenges. Ultimately, even aside from the "simpler" computational challenges, I think automated fact-checking is non-trivial and will be an ongoing challenge because it's an adversarial task: politicians will adjust to the state-of-the-art and attempt to phrase their statements in ways that evade detection. But, I also think that this adversarial nature makes the task extremely interesting.

I'm not surprised the the authors do a decent job of detecting facts from opinions (non-factual sentences). Though I want to know more about how their approach dovetails or builds upon Jan Weibe's work. Weibe's lab focuses on "subjectivity analysis" -- recognizing expressions of opinions in text. Given that her lab has dedicated significant time and energy to this area, I'd imagine that many of her tools could be used as a building block of a larger system.

I'm also not surprised that detecting unimportant facts is challenging. Aside from the obvious (there are only 400 unimportant fact sentences in the data set, so there's very little data on which to train a classifier) I imagine that "importance" is hard to infer from standard text-based features, rather than broader context.

Review: I like that the authors created a resource of crowdsourced labeled sentences, though I was slightly disappointed to find that only 8,000 of the 20,000 were retained in the end. Is this resource publicly available? I'd also be interested to know how these human-generated labels match up with labels obtained via one of Weibe's tools.

What's AlchemyAPI? What's the training data used for their sentiment analyzer? In my experience, sentiment is heavily domain-dependent. So, for example, applying a sentiment analyzer trained on data from domain X to new data from domain Y yields near-random results. Is the data used to train AlchemyAPI actually political text? If, for example, it's movie reviews or product reviews, then the output will be junk.

I'm not surprised that the cardinal number POS tag is important. But I wanted to know which other features were important? And which were not important? Also, what was the justification for using these particular features. (For example, why is sentence length likely to be an indicator of fact vs. opinion? This wasn't clear to me.)

I'm not super impressed with the machine learning methods used (very simple, off-the-shelf classifiers) but I don't think that's a problem -- one has to start somewhere with any new task and starting simple is generally a good idea. That said, I did find the presentation of results hard to follow because many different metrics were presented (e.g., precision, precision@K, nDCG) without my explanation of each. But, of course, 5 pages isn't very long :-)

Overall, I think this paper clearly relevant to the conference. I think automated checking of political facts is an interesting research area and one that isn't on most computer scientists' radars. I would like to see this work presented as a demo/poster. Thought I'd also be interested in seeing a broader panel discussion of the need for, role of, and challenges to automated fact-checking systems.

### Review 3

Review: This paper presents the CLAIMBUSTER system, which seeks to automate fact checking. Overall, the system presented in interesting and while the system is far from perfect (it is a hard problem), this seems to be a fine too. Evaluation seems limited, but the case study is interesting. Overall, the introduction and motivation in the paper is too long. We all know this is important, please provide more details on the approach / system