ClaimViz: Visual Analytics for Identifying and Verifying Factual Claims

Md Main Uddin Rony* — University of Maryland

Enamul Hoque† — York University

Naeemul Hassan‡ — University of Maryland

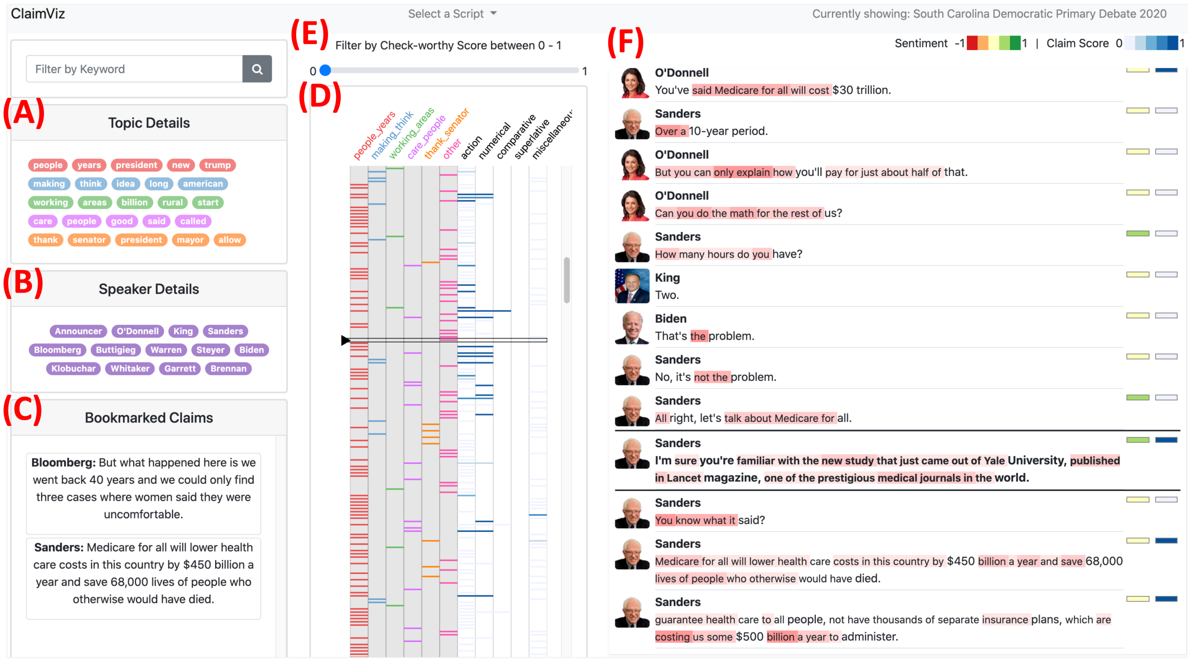

Figure 1: *ClaimViz* interface supports faceted exploration of a debate transcript based on discussion topics (A), speakers (B) and claims' check-worthiness (E). The Minimap at the middle (D) visualizes how the potentially check-worthy sentences are distributed across different topics and claim types (e.g. numerical) and allows the user to quickly locate a check-worthy claim made by a speaker in the transcript view (F). The user can bookmark any potential claims for performing the verification task (C).

Abstract

Verifying a factual claim made by public figures, aka fact-checking, is a common task of the journalists in the newsrooms. One critical challenge that fact-checkers face is- they have to swift through a large amount of text to find claims that are check-worthy. While there exist some computational methods for automating the fact-checking process, little research has been done on how a system should combine such techniques with visualizations to assist fact-checkers. *ClaimViz* is a visual analytic system that integrates natural language processing and machine learning methods with interactive visualizations to facilitate the fact-checking process. The design of *ClaimViz* is based on analyzing the requirements of real fact-checkers and our case studies demonstrate how the system can help users to effectively spot and verify claims.

Index Terms: Human-centered computing—Visualization—Visualization design and evaluation methods

1 Introduction

Fact-checking is a task in journalism where the goal is to assess the veracity of claims made by public figures; especially politicians. Although it is a common task in all newsrooms, due to the large amount of misinformation and to stay away from political biases, independent fact-checking organizations such as *PolitiFact.com*, *Factcheck.org* have started to form and flourish in recent years. According to Duke Reporters Lab, the number of fact-checking organizations in the world has increased five times since 2014 [26]. Due to the large volume of content to be checked, fact-checking has been the subject of calls from the journalism community to develop tools to automate this task [7, 10]. Researchers are working to automate the fact-checking process; mostly by combining machine learning, natural language processing and information retrieval methods to automatically spot the factual claims in text corpora (e.g. debates, interviews, tweets) and subsequently verifying the claim based on relevant information sources [14, 36, 37]. While most of these works have focused on computation methods, there remains a knowledge gap of how to effectively design for the users (e.g., fact-checkers, media analysts, journalists) to take advantage of these methods.

In this paper, we conducted a design study with fact-checkers to understand how combining natural language processing with visualization techniques could support the user in spotting and verifying claims. After identifying the challenges and requirements of these fact-checkers, we designed and evaluated *ClaimViz*, a visual analytic system, to support fact-checkers in finding and verifying factual claims. *ClaimViz* (Figure 1) supports the user to explore the potentially check-worthy claims in a large transcript of conversations such as debates and interviews through multiple facets (e.g. topics, speakers) and then select claims for further verification. Our contributions include **(1)** a visual analytic system called *ClaimViz* which makes novel integration between natural language processing and interactive visualization techniques to support fact-checkers and **(2)** the evaluation of the tool with four professional fact-checkers which provides some initial evidence of the potential utility of *ClaimViz* and provides future directions for developing visual analytic tools in the area of computational journalism.

*e-mail: mrony@umd.edu
†e-mail: enamulh@yorku.ca
‡e-mail: nhassan@umd.edu

2 Related Work

Automation in fact-checking: The concepts of automated fact-checking and misinformation detection are related yet cover two separate regions on a common information disorder spectrum [38]. While misinformation detection usually focuses on assessing credibility of articles, automated fact-checking focuses on assisting journalists in vetting and verifying factual statements [10]. Hassan et al. primarily delineated the vision of automating the fact-checking process and identified the computational and journalistic challenges that are involved in this process [11]. One of the critical computational challenges, assessing check-worthiness of claims, has been addressed by researchers using machine learning models [13,21,31]. Another challenge- verification, has been addressed by computation methods that can be divided into two categories by and large- *AI based* and *Database based*. The AI based systems primarily focus on the patterns and features from the data to predict the veracity using supervised learning [18,23,40]. While these methods are good at predicting veracity for new claims, the assumption of existence of latent patterns to differentiate fake news from real news may not be true in carefully curated disinformation contents. The Database based approaches assume that knowledge bases have enough relevant information to check claims [6]. While these approaches may explain the result better than the AI based solutions, these methods are incapable of assessing fresh claims. ClaimBuster [12, 14] presented the first end-to-end prototype of an automated fact-checking system that uses the Database approach. It identifies the check-worthy claims and then predicts the veracity of claims by querying knowledge bases such as Wolfram Alpha and Freebase. Ciampaglia et al. [6] uses Wikipedia as a knowledge network and uses path enumeration between entities to assess the veracity of a claim. There are some works which combine both AI and Database approaches. For instance, Shi et al. [36] leverage knowledge graphs for retrieving information about entities in a claim and predict links between entities to quantify the truthiness of the claim. Others attempted to fact-check visual images instead of text. For instance, DejaVi supports journalists in identifying misinformation from social media images [27] by allowing them to flag images collaboratively, so that others can find those images as well as their near-duplicate images.

Visualizations for fact-checking: Although there has been much research on computationally modeling the fact-checking process, limited work has been done on the integration of these automated solutions in the newsrooms among the fact-checkers. Existing systems like ClaimBuster [14] and ClaimPortal [24] only provide limited interactive features like filtering based on claim score and do not provide adequate visualization support to spot and verify claims. Nguyen et al. [30] present a mixed-initiative approach for presenting the fact-checking predictions to help users understand how the underlying model arrived at its prediction. Again, their interface provides some filtering options to help users interact with the features of the model but it does not provide a visual overview of the fact-check worthy claims. Kurdani et al. [19,20] present a visual analytic system to support the identification of misinformation in social media through analysis of text, social network, image, and language feature. In this paper, we closely work with the professional fact-checkers and design, *ClaimViz*, a visual analytic system for assisting the fact-checking process by leveraging natural language processing and visualization methods. We focus on fact-checking instead of misinformation detection and deal with speech transcripts (e.g. meetings, interviews, debates) as opposed to social media texts which are often constrained by character limits.

Visual text analysis for multi-party conversations: There has been growing interest in combining automatic text analysis and interactive visualizations to support the exploration of multi-party conversations. Researchers developed interactive visualizations to show automatically extracted topics and sentiments from conversations [8, 15, 17] as well as to show conversational dynamics using temporal episodes [34]. There have also been attempts to visualize argumentation patterns to support analysis of a debate's deliberative content [8, 9, 32]. While the above body of work are helpful in understanding multi-party conversations they are not designed to deal with specific journalistic tasks of fact-checking.

3 The Fact-checking domain

3.1 Requirement Gathering

In order to design ClaimViz, we followed visualization design study methodologies [29,33] where we started by interviewing three professional fact-checkers from *Politifact.com* and *Duke Reporters Lab*. The interview was open-ended in nature and was run in a focus group setting so that it could open up a diverse range of issues and possible solutions. Before the interview, we demonstrated the early version of *ClaimViz* (which was based on Claimbuster [14], a system that all three fact-checkers were familiar with) so that the interviewees get a sense of the intended goal. Following that, we gathered feedback on the early mockup and further requirements.

The interviewees provided important insights into their current work practices and challenges. They mentioned that they primarily focus on spoken texts where people are often forced to express what they have in mind as opposed to social media where the texts are short and speakers are usually more careful in what they express. They have tested ClaimBuster in their claim identification process in a limited scope. The workflow started by ClaimBuster monitoring news sources that are predefined by the organizations. It identified check-worthy claims from these sources and prepared an email newsletter with highly check-worthy claims and sent the letter to all fact-checkers around the country. While they felt that without ClaimBuster-like tool it would be very difficult to constantly monitor the new sources for spotting claims they also pointed out several tasks that are not currently well supported such as the lack of ability for human to provide input and absence of context information [3].

3.2 Tasks Analysis

Our conversation with domain experts revealed several important analytical tasks as follows:

T1: Spot and read claims in context: The fact-checkers want to get a visual overview of how check-worthy claims are distributed across a transcript so that they can quickly spot the interesting claims. At the same time, just presenting the sentence containing the claim is not enough, they need the ability to read a few sentences before or after the claim to understand the proper context of the claim.

T2: Understand topics of factual claims: Since a transcript may involve several topics of discussion, interviewees wanted to understand how claims are distributed across different topics and when topic transitions occur. They also wanted topics represented with meaningful short-phrases and the ability to merge similar topics together whenever necessary.

T3: Find speaker attribution: Fact-checkers want to know who made the claim under examination. Conversely, they may want to find all claims made by a specific speaker of interest.

T4: Annotating claims: Fact-checkers need to mark down claims and record them for further verification. They also want to forward the selected claims to others, preferably through emails.

T5: Find relevant evidence for a given claim: After spotting claims in a transcript, fact-checkers must analyze external sources and knowledge-base to verify the truthfulness of the claims. For achieving this goal, they need to find all the relevant sources that may provide supporting or opposing evidence to the claim.

3.3 Text Analytics

Our system applies several text analysis techniques on the transcribed document and then visualizes the results to the user (Figure 2). To support the fact-checking tasks, the system predicts: (1) the check-worthiness and (2) the claim type. We also applied techniques for topic and sentiment analysis as well as evidence mining.

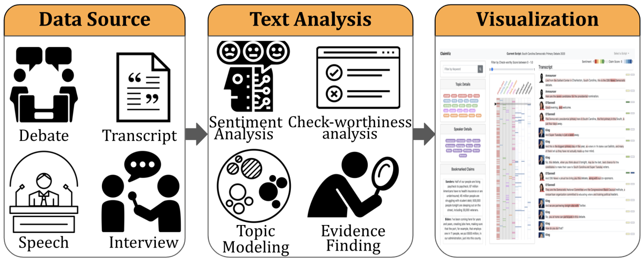

Figure 2: System architecture of the *ClaimViz* system

Claim check-worthiness prediction: We adopted a Bidirectional Long Short Term Memory Networks (LSTM) model to predict the check-worthiness of each sentence [35]. The system converts the input sentences into vectors by applying *word2vec* [28] and then feeds them to the model. We chose a bidirectional LSTM because it captures long dependencies within the input sequence from both forward and backward directions. The model also uses the attention mechanism to capture certain input time steps that the model should provide more focus on. More importantly, since this mechanism returns the weight of each token indicating the amount of attention given to the token while predicting the check-worthiness score, it enables us to visualize how much a word contributes to the score.

We used the ClaimBuster dataset to train and evaluate our model [4]. This dataset contains 23,533 sentences extracted from U.S. general election presidential debates and human annotators categorized them into three groups (Check-worthy Factual Sentence, Unimportant Factual Sentence, and Non-factual Sentence). We performed 5-fold cross-validation to evaluate the model with respect to precision, and recall. Our model achieved good performance with 75% micro-averaged precision and 75% micro-averaged recall.

Claim type classification: We devised a rule-based classification technique that categorizes a claim into one of the 5 types: (1) Action (e.g. "My opponent supported Iraq war"), (2) Numerical: "Pete has gotten funding from over 50 billionaires", (3) Comparison: (e.g. "I've run the city which is almost the same size – bigger than most countries in the world."), (4) Superlative: (e.g. "I was the mayor of the largest, most populous city in the United States"). The classifier first applies Stanford CoreNLP to detect POS tags and basic dependencies between the tokens [25]. Then, it applies a set of heuristics to identify the claim types. For example, the presence of POS tag `JJS` (Adjective, superlative) in the sentence usually suggests that the claim type is *Superlative* while `JJR` (Adjective, comparative) can be indicative of a *Comparative* type of claim. Similarly, we checked the presence of the POS tag `CD` (Cardinal Number) to identify a numerical type of claim. Along with the POS tags we also considered the dependencies among the tokens in some heuristics. For example, to identify the action type of claim, we checked if the sentence is in *Subject-Verb-Object* format. If a claim does not get assigned to any of the four types, we put that in 'miscellaneous' type.

Topics and sentiment mining: Instead of using the traditional topic modeling algorithms (e.g., Latent Dirichlet Allocation [5], Latent Semantic Analysis [22]) which may suffer from the high word co-occurrence patterns presented in shorter text, we used Biterm Topic Modeling (BTM) [39]. BTM generates the topics by addressing the aggregated word co-occurrence patterns directly. We considered each sentence of the transcript as an individual document and applied BTM on it to generate 5 topics. If a sentence does not have a minimum probability to be assigned to any topic, we assigned it to the 'other' topic. We generated 5 most representative words from each topic based on word frequency analysis and TF-IDF measure. For sentiment analysis, we used textblob [1], a python package which gives a sentiment score between -1 to +1 for each sentence.

Evidence mining: In order to help users to verify a claim, we used Google Search Library [2] that retrieves relevant documents given a claim sentence. We then measured the similarity between the

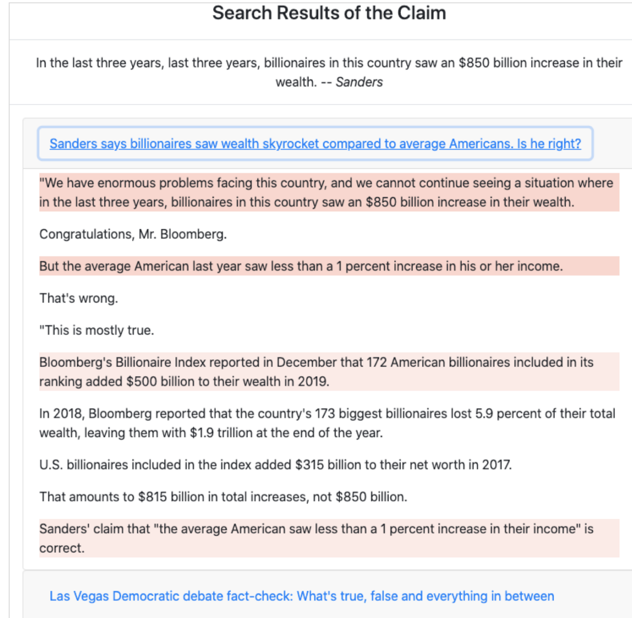

Figure 3: As the user clicks on the search button beside a sentence in the Transcript View, the system retrieves related articles from the Web and then highlights sentences that are relevant to the claim sentence.

claim sentence and each sentence of the retrieved documents. For this similarity measure, we first represented input sentence pairs using *word2vec* [28] and then calculated the cosine similarity between the two sentences to find out which sentences from the retrieved documents are the most similar to the given claim.

4 *ClaimViz* Visualization Design

4.1 Design Goals

Based on our user requirements gathering and tasks analysis, we derive the following design goals:

DG 1 Overview and filter: The system should present a visual overview of check-worthy claims to help the user in spotting claims and reading them from a long transcript (**T1**). Subsequently, users can choose highly check-worthy claims by filtering based on claim scores and save them for future verification (**T4**).

DG 2. Support faceted exploration: The user should be able to explore claims made by a speaker about a specific discussion topic (**T2, T3**) through faceted exploration.

DG 3. Enhance model transparency: The system should convey why it predicts a sentence as claim check-worthy to help the user spotting the claim more effectively (**T1, T4**).

DG 4. Show relevant evidence on demand: After finding a claim, the user should be able to see relevant information evidence from Web on demand for verifying that claim (**T5**).

4.2 Interaction and Visual Design

ClaimViz uses multiple linked views (see Figure 1) to help users performing the fact-checking task. The interface shows a high-level overview of claims, a transcript view as well as the topics and speakers involved in a conversation. We now describe and justify most of the basic visual encodings and interactions used in *ClaimViz*.

The **Minimap** shows a visual overview of the whole transcript (Figure 1D) so that the fact-checker can locate check-worthy claims about specific topics (**DG1**). It consists of two groups of columns; one group indicates to which topic each sentence, represented by a thin line, belongs to and the other group indicates to which claim type each sentence belongs to. Lines within each column are vertically arranged based on their positions in the transcripts while their fill-color indicates their claim scores or topics they belong to. Users can filter sentences in the Minimap by claim scores using the slider to find the ones that are highly check-worthy. After locating a sentence that looks like a highly check-worthy claim the user can click on the line, which results in highlighting the actual sentence in the Transcript View so that she can read that sentence in context.

The **Transcript View** shows the raw text (Figure 1F) in a scrollable pane along with the claim score and the sentiment score for each sentence. To help the user understand why the system predicts a sentence as check-worthy (**DG3**), it highlights the words with a dark tone if they contribute more towards high claim scores according to our attention based model for check-worthiness prediction. As the user scrolls through the transcript, the system moves a cursor (encoded as a black rectangle) in the Minimap View accordingly.

The user can perform faceted explorations based on the lists of topics and speakers **(DG2)** (Figure 1A, B), left). Selecting a keyword within a topic results in highlighting the corresponding sentences where that keyword appears, both within the Minimap and Transcript View. The user can also search for a specific keyword through a textbox.After examining a sentence, if the user feels that this is check-worthy, she can add it to the bookmark panel by clicking on the bookmark icon. Also, if the user clicks on the search icon beside a sentence, the system retrieves relevant documents from the web that may provide evidence about the claim **(DG4)**. The relevant documents are shown in a popup panel where sentences related to the claim are highlighted to help users verify the claim (Figure 3).

5 Expert Case Studies

To test the efficacy of *ClaimViz*, we ran case studies with four domain experts on two different transcripts- the first democratic primary debate and the first GOP primary debate of the 2016 US presidential election. Both transcripts are similar in length (1530 and 1378 sentences, respectively). Through the studies we wanted to understand: i) whether *ClaimViz* helps target users in finding and verifying claims effectively. ii) which visualization features worked and did not work? and iii) what can we learn from their feedback to improve future development of factual claim checking tools?

Participants: We ran the study with four domain experts (3 male, 1 female, age range 18-40 years). The participants have extensive amount of fact-checking experience (1-4 years). They also confirmed that they perform fact checking frequently (several times a week). Most of them chose the topics of politics as their first choice of interests followed by health, science, and religion.

Procedure: After filling up the pre-study questionnaire, the participant was provided with a brief tutorial about how *ClaimViz* works. Then the participant performed two types of tasks: i) four target criteria tasks where they would have to find check-worthy claims about some specific topics or speakers or both (e.g., "Find claims related to statistics about work hours and wages made by Sanders."). ii) an exploratory task with a different transcript where they would have to find five most check-worthy claims according to their own interests and they were free to use as much time as they needed. Participants were requested to record their responses by bookmarking the claims they found. The participant also filled up a post-study questionnaire and went through a brief semi-structured exit interview.

Results: According to the post-study questionnaire, all participants found the *ClaimViz* interface to be effective in identifying check-worthy claims. In particular, they mostly agreed that showing the topic and claim score of each sentence in the Minimap is useful (3 out of 4). Similarly, they found that filtering by claim scores in the Minimap was useful to locate check-worthy claims (3 out of 4). All participants liked the interactive features of faceted exploration by topics and speakers while 3 participants found that highlighting words that are indicative of check-worthy claims is very useful.

We examined the claims picked by the participants using *ClaimViz* during their tasks by checking if any of their claims were also chosen by CNN, Washington Post, FactCheck.org, or PolitiFact fact-checkers as their top picks. We found that all participants chose at least one claim that was also chosen by one of these organizations. We also observed that among 3 of the 4 target criteria tasks, at least one participant picked a claim that was also fact-checked by these four organizations. For instance, when the participants were asked to *find statistical claims by Bernie Sanders about wages*, one participant picked 3 claims out of which 2 were selected and fact-checked by one of the four organizations. We argue that the results are encouraging because the transcript was very long with 1530 sentences and even though there were many highly check-worthy claims (276 claims had check-worthiness score $>= 0.5$), the participants picked 5 of the 30 claims that were verified by the four organizations.

We also analyzed the interviews to understand the overall reactions regarding *ClaimViz* and suggestions for improvement. Overall, participants were quite impressed with the tool. According to P3 *"...Amazing! Even a small fact-checking task could be painful as I need to go through a huge volume of content. But using this tool I will take much less this time for finding the claims"*. P2 suggested that *"The tool is surprisingly helpful... It could not only be very effective for fact-checking but also for other journalistic tasks on debate transcripts where we need to find opinions from speakers"*.

Post-study interviews also revealed that while *ClaimViz* mostly met the identified task requirements there were some concerns from participants. P1 was concerned that people who are not familiar with technology may take some time to understand how it works. Moreover, while all participants agreed that the filtering by claim score was helpful, two of them did not find the word highlighting feature equally helpful *"When I selected the high claim scores through the slider, I quickly managed to find check-worthy claims, but when I looked at colored words not all of them are indicative of check-worthy claims"* (P4). P3 suggested that *"While fact-checking I read the whole sentence anyway, so highlighting words in transcripts was not that helpful to me"*. A suggestion was to provide additional way to control the threshold of word weights so that only the words that are highly indicative of check-worthy claims can be highlighted. P3 also suggested that while fact-checking, the sentiment score is not as helpful as the claim score so showing it in the interface is not so relevant. Finally, regarding the faceted exploration feature, P4 pointed out that while selecting sentences by a speaker was useful, selecting topic keywords was not always useful and sometimes misleading *"I selected the 'policy' keyword hoping that the related sentences will talk about foreign policies but the selected sentences were not really about that"*. This could be attributed to automatic topic modeling which is known to be noisy and inaccurate especially for conversational text [16]. Some participants suggested additional features to enhance the fact-checking process. P2 suggested that showing the overall sentiment distributions for each speaker could be useful. P1 suggested that the tool only supports checking textual content, but in future it would be more helpful to support fact-checking on images.

6 Conclusion and Future Work

In this paper, we present *ClaimViz*, a visual analytic system which tightly integrates natural language processing, machine learning, and information visualization techniques to support fact-checkers in the newsroom. Our system supports the user to interactively filter sentences in a long transcript to find the highly check-worthy claims followed by verifying the claims by finding potential evidence that may support or oppose the claim. To the best of our knowledge, *ClaimViz* is the first visual analytic system that is designed to empower fact-checkers to effectively verify claims from spoken texts (e.g. debates, interviews). Our case studies with four domain experts suggest the usefulness of *ClaimViz* in helping the user to spot and verify claims from long transcripts. The *ClaimViz* system is available at http://claimviz.umd.edu/. In the future, we will extend our system for supporting fact-checking with other corpora (e.g. social media texts). We also plan to enrich the verification component with further linguistic analysis such as stance detection and argument mining.

References

[1] Textblob, 2020. https://textblob.readthedocs.io/en/dev/.

[2] Welcome to googlesearch's documentation, 2020. https://python-googlesearch.readthedocs.io.

[3] B. Adair, M. Stencel, C. Clabby, and C. Li. The human touch in automated fact-checking. In *Proceedings of the Computation+ Journalism Symposium*, 2019.

[4] F. Arslan, N. Hassan, C. Li, and M. Tremayne. A benchmark dataset of check-worthy factual claims. In *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 14, pp. 821–829, 2020.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[6] G. L. Ciampaglia, P. Shiralkar, L. M. Rocha, J. Bollen, F. Menczer, and A. Flammini. Computational fact checking from knowledge networks. *PloS one*, 10(6):e0128193, 2015.

[7] S. Cohen, J. T. Hamilton, and F. Turner. Computational journalism. *Communications of the ACM*, 54(10):66–71, 2011.

[8] M. El-Assady, V. Gold, C. Acevedo, C. Collins, and D. Keim. Contovi: Multi-party conversation exploration using topic-space views. In *Computer Graphics Forum*, vol. 35, pp. 431–440. Wiley Online Library, 2016.

[9] M. El-Assady, A. Hautli-Janisz, V. Gold, M. Butt, K. Holzinger, and D. Keim. Interactive visual analysis of transcribed multi-party discourse. In *Proceedings of ACL 2017, System Demonstrations*, pp. 49–54. Association for Computational Linguistics, July 2017.

[10] D. Graves. Understanding the promise and limits of automated fact-checking. Technical report, Reuters Institute for the Study of Journalism, 2018.

[11] N. Hassan, B. Adair, J. T. Hamilton, C. Li, M. Tremayne, J. Yang, and C. Yu. The quest to automate fact-checking. In *The Computation+Journalism Symposium*, 2015.

[12] N. Hassan, F. Arslan, C. Li, and M. Tremayne. Toward automated fact-checking: Detecting check-worthy factual claims by claimbuster. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1803–1812. ACM, 2017.

[13] N. Hassan, C. Li, and M. Tremayne. Detecting check-worthy factual claims in presidential debates. In *Proceedings of the 24th acm international on conference on information and knowledge management*, pp. 1835–1838. ACM, 2015.

[14] N. Hassan, G. Zhang, F. Arslan, J. Caraballo, D. Jimenez, S. Gawsane, S. Hasan, M. Joseph, A. Kulkarni, A. K. Nayak, et al. Claimbuster: The first-ever end-to-end fact-checking system. *Proceedings of the VLDB Endowment*, 10(12):1945–1948, 2017.

[15] E. Hoque and G. Carenini. Convis: A visual text analytic system for exploring blog conversations. In *Computer Graphics Forum*, vol. 33, pp. 221–230. Wiley Online Library, 2014.

[16] E. Hoque and G. Carenini. Interactive topic modeling for exploring asynchronous online conversations: Design and evaluation of convisit. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 6(1):1–24, 2016.

[17] E. Hoque and G. Carenini. MultiConVis: A visual text analytics system for exploring a collection of online conversations. In *Proceedings of the ACM Conference on Intelligent User Interfaces (IUI)*, pp. 96–107, 2016.

[18] Z. Jin, J. Cao, Y. Zhang, and J. Luo. News verification by exploiting conflicting social viewpoints in microblogs. In *Thirtieth AAAI Conference on Artificial Intelligence*, pp. 2972–2978, 2016.

[19] A. Karduni, I. Cho, R. Wesslen, S. Santhanam, S. Volkova, D. L. Arendt, S. Shaikh, and W. Dou. Vulnerable to misinformation?: Verifi! In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, IUI '19, pp. 312–323, 2019.

[20] A. Karduni, R. Wesslen, S. Santhanam, I. Cho, S. Volkova, D. Arendt, S. Shaikh, and W. Dou. Can you verifi this? studying uncertainty and decision-making about misinformation using visual analytics. In *Proceedings of the International AAAI Conference on Web and Social Media*, pp. 151–160, 2018.

[21] L. Konstantinovskiy, O. Price, M. Babakar, and A. Zubiaga. Towards automated factchecking: developing an annotation schema and benchmark for consistent automated claim detection. *arXiv preprint arXiv:1809.08193*, 2018.

[22] T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch. *Handbook of latent semantic analysis*. Psychology Press, 2013.

[23] J. Ma, W. Gao, and K.-F. Wong. Detect rumors in microblog posts using propagation structure via kernel learning. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 708–717, 2017.

[24] S. Majithia, F. Arslan, S. Lubal, D. Jimenez, P. Arora, J. Caraballo, and C. Li. ClaimPortal: Integrated monitoring, searching, checking, and analytics of factual claims on twitter. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 153–158. Association for Computational Linguistics, Florence, Italy, July 2019.

[25] C. D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pp. 55–60, 2014.

[26] J. L. Mark Stencel. Reporters' Lab fact-checking tally tops 200, 2019. https://reporterslab.org/reporters-lab-fact-checking-tally-tops-200.

[27] H. Matatov, A. Bechhofer, L. Aroyo, O. Amir, and M. Naaman. Dejavu: A system for journalists to collaboratively address visual misinformation. In *Computation+ Journalism Symposium. Miami, FL. https://scholar. harvard. edu/files/oamir/files/dejavu-system-journalists_3. pdf*, 2018.

[28] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv:1301.3781*, 2013.

[29] T. Munzner. A nested model for visualization design and validation. *IEEE Trans. Visualization & Computer Graphics*, 15(6):921–928, 2009.

[30] A. T. Nguyen, A. Kharosekar, S. Krishnan, S. Krishnan, E. Tate, B. C. Wallace, and M. Lease. Believe it or not: Designing a human-ai partnership for mixed-initiative fact-checking. In *The 31st Annual ACM Symposium on User Interface Software and Technology*, pp. 189–199. ACM, 2018.

[31] A. Patwari, D. Goldwasser, and S. Bagchi. Tathya: A multi-classifier system for detecting check-worthy statements in political debates. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pp. 2259–2262, 2017.

[32] B. Plüss, F. Sperrle, V. Gold, M. El-Assady, A. Hautli, K. Budzynska, and C. Reed. Augmenting public deliberations through stream argument analytics and visualisations. In *Leipzig Symposium on Visualization in Applications*, 2018.

[33] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: reflections from the trenches and the stacks. *IEEE Trans. Visualization & Computer Graphics*, 18(12):2431–2440, 2012.

[34] D. Seebacher, M. T. Fischer, R. Sevastjanova, D. A. Keim, and M. El-Assady. Visual Analytics of Conversational Dynamics. In T. v. Landesberger and C. Turkay, eds., *EuroVis Workshop on Visual Analytics (EuroVA)*. The Eurographics Association, 2019. doi: 10.2312/eurova.20191130

[35] M. Seo, A. Kembhavi, A. Farhadi, and H. Hajishirzi. Bidirectional attention flow for machine comprehension. *arXiv preprint arXiv:1611.01603*, 2016.

[36] B. Shi and T. Weninger. Discriminative predicate path mining for fact checking in knowledge graphs. *Knowledge-based systems*, 104:123–133, 2016.

[37] J. Thorne and A. Vlachos. Automated fact checking: Task formulations, methods and future directions. *arXiv preprint arXiv:1806.07687*, 2018.

[38] C. Wardle. Fake news. it's complicated. 2017.

[39] X. Yan, J. Guo, Y. Lan, and X. Cheng. A biterm topic model for short texts. In *Proceedings of the 22nd international conference on World Wide Web*, pp. 1445–1456, 2013.

[40] A. Zubiaga, M. Liakata, and R. Procter. Exploiting context for rumour detection in social media. In *International Conference on Social Informatics*, pp. 109–123. Springer, 2017.