

# A News Verification Browser for the Detection of Clickbait, Satire, and Falsified News

Victoria L. Rubin<sup>1</sup>, Chris Brogly<sup>1</sup>, Nadia Conroy<sup>1, 3</sup>, Yimin Chen<sup>1</sup>, Sarah E. Cornwell<sup>1</sup>, and Toluwase V. Asubiaro<sup>1, 2</sup>

**1** Language and Information Technology Research Lab (LiT.RL), Faculty of Information and Media Studies, The University of Western Ontario, London, Ontario, Canada. **2** E. Latunde Odeku Medical Library, College of Medicine, University of Ibadan, Ibadan, Nigeria **3** Media Sonar Technologies, London, Ontario, Canada.

DOI: [10.21105/joss.01208](https://doi.org/10.21105/joss.01208)

## Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 12 December 2018

Published: 09 March 2019

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC-BY).

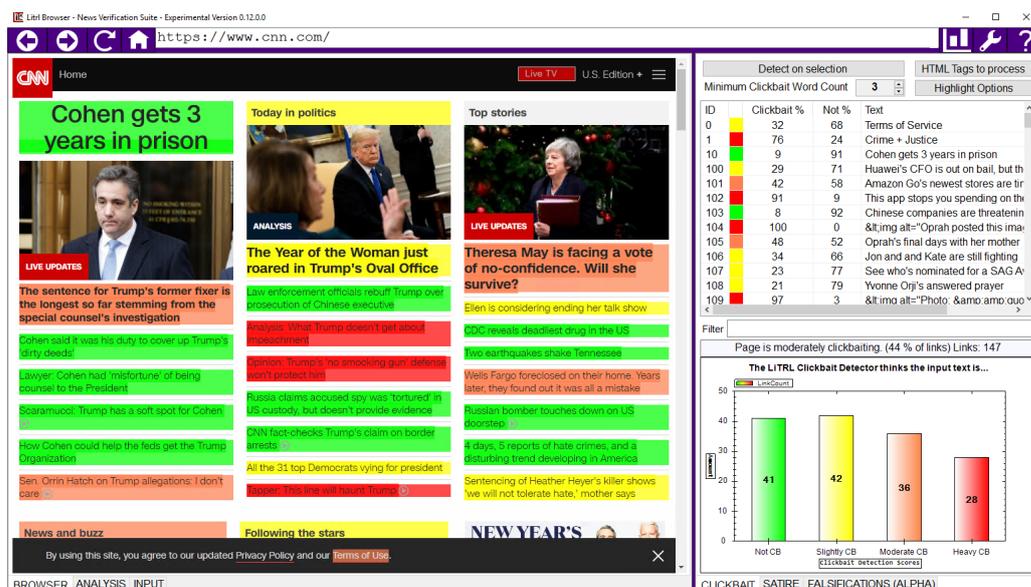
## Description-in-Brief

The LiT.RL News Verification Browser is a research tool for news readers, journalists, editors or information professionals. The tool analyzes the language used in digital news web pages to determine if they are clickbait, satirical news, or falsified news, and visualizes the results by highlighting content in color-coded categories. Although the clickbait, satire, and falsification detectors perform to certain accuracy levels on test data, during real-world internet use accuracy may vary. The browser is not a replacement for digital literacy and is not always correct. All processing is completed on the local machine - results are not sent to or from a remote server. Results may be saved locally to a standard SQLite database for further analysis.

## Summary

Widespread adoption of internet technologies has changed the way that news is created and consumed. The current online news environment is one that incentivizes speed and spectacle in reporting at the cost of fact-checking and verification, encouraging the proliferation of misinformation and disinformation. The LiT.RL News Verification (NV) Browser is a system that offers a first step counter-measure by automatically detecting and highlighting clickbait (to 94% accuracy on a test set of 5670 texts), satirical fake news (to 84% accuracy on a test set of 95 texts), and fabricated news (to 71% accuracy on a test set of 28 texts). The browser was built to study the effectiveness of these deception detectors when applied to online news websites, where the accuracy of these detectors may vary considerably given the variety of styles, formats and content of online news. The system is presented to users as a set of assistive technologies built into a stand-alone browser tuned to identify different varieties of fakes (Rubin, 2017; Rubin, Chen, & Conroy, 2015). Our algorithms look for patterns of subtle lexico-syntactic features in text. Images, audio and video formats are unsupported. The core of the functionality is in the natural language processing (NLP) of textual data and automated classification of results with machine learning using support vector machines.

The LiT.RL NV Browser offers three discrete functionalities: a. Detection of clickbait headlines (Brogly & Rubin, 2019; Chen, Conroy, & Rubin, 2015a); b. Detection of satirical article content (Rubin, Conroy, Chen, & Cornwell, 2016); c. Detection of falsified news articles (Asubiaro & Rubin, 2018, 2019; Rubin & Conroy, 2012).



**Figure 1:** Figure 1: Screenshot of LiT.RL News Verification Browser clickbait detection on the CNN homepage (December 12, 2018)

Each is implemented as a separate overlay option through three tabs at the bottom right of the browser. Currently in its proof of concept stage, the system runs over a batch of website content, such as news feed on a website, and visualizes the results by highlighting content in red, orange or green, by analogy with the traffic stop-light.

## Limitations

This browser is meant to augment our human discernment, rather than replace it, by highlighting potentially false information which may require further scrutiny. Digital literacy is key for everyone to effectively evaluate potential misinformation online, and the LiT.RL Browser is NOT a replacement for that (Chen, Conroy, & Rubin, 2015b). During real-world internet use, the accuracy of the detectors in this tool may vary. News readers' critical thinking remains key to navigating the increasingly fraught online information landscape. The News Verification Browser is a research tool - not a replacement for a day-to-day web browser. It should not be used where security is critical.

## Availability

The source code is openly available on GitHub (Rubin et al., 2018) under the GPLv3 license for anyone in the research and development community to use or improve on. The public can download the browser for experimentation on their own computers with "no strings attached".

## Acknowledgements

This research has been funded by the Government of Canada Social Sciences and Humanities Research Council (SSHRC) Insight Grant (#435-2015-0065) awarded to Dr. Victoria

L. Rubin for the 2015-2018 project entitled “Digital Deception Detection: Identifying Deliberate Misinformation in Online News”.

## References

- Asubiaro, T., & Rubin, V. L. (2018). Comparing features of fabricated and legitimate political news in digital environments (2016-2017). In *Proceedings of the 81st ASIST annual meeting* (pp. 747–750). Vancouver, Canada. doi:[10.1002/pr2.2018.14505501100](https://doi.org/10.1002/pr2.2018.14505501100)
- Asubiaro, T., & Rubin, V. L. (2019). Falsified news detection methodology. *Unpublished article*.
- Brogly, C., & Rubin, V. L. (2019). Detecting clickbait: Here’s how to do it. *Unpublished article under review*.
- Chen, Y., Conroy, N. J., & Rubin, V. L. (2015a). Misleading online content: Recognizing clickbait as “false news”. In *Proceedings of the 2015 ACM workshop on multimodal deception detection, WMDD ’15* (pp. 15–19). New York, NY, USA: ACM. doi:[10.1145/2823465.2823467](https://doi.org/10.1145/2823465.2823467)
- Chen, Y., Conroy, N. J., & Rubin, V. L. (2015b). News in an online world: The need for an “automatic crap detector”. In *Proceedings of the 78th ASIST annual meeting: Information science with impact: Research in and for the community, ASIST ’15* (pp. 81:1–81:4). Silver Springs, MD, USA: American Society for Information Science. doi:[10.1002/pr2.2015.145052010081](https://doi.org/10.1002/pr2.2015.145052010081)
- Rubin, V. L. (2017). News verification suite: Towards system design to supplement reporters’ and editors’ judgements. In *Proceedings of the 45th annual conference of the canadian association for information science/l’Association canadienne des sciences de l’information (cais/acsi2017)*, CAIS/acsi 2017. Edmonton, Alberta: Cameron Library, University of Alberta. Retrieved from <https://journals.library.ualberta.ca/ojs.cais-acsi.ca/index.php/cais-ascii/article/view/1030>
- Rubin, V. L., & Conroy, N. (2012). Discerning truth from deception: Human judgments and automation efforts. *First Monday*, 17(3). doi:[10.5210/fm.v17i3.3933](https://doi.org/10.5210/fm.v17i3.3933)
- Rubin, V. L., Brogly, C., Conroy, N., Chen, Y., Cornwell, S. E., & Asubiaro, T. (2018). [https://github.com/litrl/litrl\\_code](https://github.com/litrl/litrl_code). doi:[10.5281/zenodo.2016628](https://doi.org/10.5281/zenodo.2016628)
- Rubin, V. L., Chen, Y., & Conroy, N. (2015). Deception detection for news: Three types of fakes. In *Proceedings of the 78th ASIST annual meeting: Information science with impact: Research in and for the community, ASIST ’15* (pp. 83:1–83:4). Silver Springs, MD, USA: American Society for Information Science. doi:[10.1002/pr2.2015.145052010083](https://doi.org/10.1002/pr2.2015.145052010083)
- Rubin, V. L., Conroy, N., Chen, Y., & Cornwell, S. E. (2016). Fake news or truth? Using satirical cues to detect potentially misleading news. In *Proceedings of the second workshop on computational approaches to deception detection* (pp. 7–17). San Diego, California: Association for Computational Linguistics. doi:[10.18653/v1/W16-0802](https://doi.org/10.18653/v1/W16-0802)