News Verification Suite:
Towards System Design to Supplement Reporters' and Editors' Judgements

Victoria L. Rubin
Language and Information Technology Research Lab (LIT.RL)
Faculty of Information and Media Studies
University of Western Ontario
FIMS & Nursing Building, Room 4046
London, Ontario, Canada N6A 5B7
vrubin@uwo.ca

Abstract

The News Verification Suite aims to provide users with a set of functions to verify information in the news. This paper offers a conceptual basis and a vision of system elements towards automated fact-checking in news production, curation, and consumption. The traditional model of journalism is compared to 'news sharing a.s.a.p.', highlighting similarities between journalistic criteria of excellence and LIS mandates for credibility and information quality. Potential steps for intervention with text-analytical technologies are identified – deception detection, rumor busting, satire labeling; they are nascent but feasible. Automated news verification can support and supplement news producers' and news readers' decision-making.

1. Introduction

1.1. The "Fake News" Problem

Current online news environments abound with deceptive information, indiscriminately labelled as "fake news". A more accurate term used for an intentional control of information to create a false belief or false conclusion in digital environments is "digital deception" (Hancock, 2012). Digital deception in the news can be of several types: rumors, hoaxes, clickbait (Chen, Conroy, & Rubin, 2015a), fraud (Rubin, 2010; Rubin, Chen, & Conroy, 2015; Rubin, Conroy, & Chen, 2015), misleading satire (Rubin, Conroy, Chen, & Cornwell, 2016), and native ads (Cornwell & Rubin, 2017). Some varieties are unique to social media, and the scale of propagation and audience reach may be unprecedented as of 2017, but the broader issues of credibility (e.g., (Rieh, 2010)), misinformation (e.g., Fox (1983)) and quality of information (e.g., Jarke and Vassiliou (1997)) have been addressed in Library and Information Science (LIS) for decades. More specifically, in the past 10 years, digital deception has been extensively studied in Natural Language Processing (NLP) and Machine Learning (ML), in the field of deception detection at the intersection of LIS with computer science, interpersonal psychology and computer-mediated communication.

Algorithmic solutions to the "fake news" problem are often inaccurately discounted as non-existent. Few LIS professionals or journalists have heard of advances in automated deception detection and rumor busting, while these techniques have made big strides since 2008 (Rubin, 2017). For instance, we were able to automate deception detection of deliberately deceptive (fabricated) news pieces and achieved 63% accuracy, which superseded an average human ability to detect a lie (Victoria L. Rubin et al., 2015). Satirical fakes (also misleading when decontextualized) can be identified automatically with even higher success rates. The Satire Detector achieved 90% precision and 84% recall on identification of the Onion- or Beaverton-style satirical news (Rubin et al., 2016). The system and the associated dataset are available for the general public[1].

Few verification mechanisms are currently available to news readers, nor are they integrated into newsroom production tools or browsers. However, automated deception detection techniques have been proven feasible since around 2008, and will likely – with time and further research and development – complement and enhance the notoriously poor human abilities to discern truth from deception, both in the news consumption and production.

[1] The Satire Detector tool is online at: http://satiredetector.fims.uwo.ca/. The associated dataset is publically available at: http://victoriarubin.fims.uwo.ca/news-verification/access-s-n-l-news-db2015-2016/.

1.2. News Production

In the traditional journalism paradigm the task of news verification is rather complex. The news production cycle involves reporters verifying what the source tells them, reporting objectively what is known to be true, separating facts from opinions, reporting on how information is obtained, and ideally submitting a copy of the news report to the fact-checkers. Editor have the discretion of vetting the information and selecting which story is newsworthy to be reported.

Detecting whether the reporter truthfully conveys the information to the audience is just one piece of the puzzle, albeit an important one. In this work I take a broader look at the news verification as a process. For instance, communications channels between the reporters and their sources should not be overlooked, as reporters may also receive inaccurate, incomplete, deceitful information and form a false conclusion. As fact-checking efforts within news organizations are being notoriously outsourced to verification by external organizations (e.g., Snopes.com and Politifact.com), there is a greater need for in-house tools and applications for newsreaders to assist them with more effective and efficient verification of information on a variety of broad information quality assessment criteria.

1.3. Problem Statement

The way reporters are producing and delivering news is undoubtedly changing (Chen, Conroy, & Rubin, 2015b), especially with the US Administration imposing new norms and legitimizing 'alternative facts' ("Donald trump's alarming alternative facts - the boston globe," 2017). The literature in journalistic instruction of analytical kind, such as best practices guide (e.g., Kovach and Rosenstiel (2010)) and technology analysts (e.g., "The Bittersweet Sweepstakes to Build an AI That Destroys Fake News," 2016) claim that there is no simple mathematic formula to the overall journalistic process of collecting, vetting and reporting information. No doubt the claim is right: the human quality of work is hard to approximate. Yet, with the scale and varieties of digital deception, trained journalists can only verify a drop in the ocean of information, and the newsreaders are ill equipped or pressed for time to do in-depth verification themselves. This situation requires a comprehensive system to support the work of professional content providers for the ultimate benefit of information users and transparency of reporting.

1.4. Research Objectives

This research aims to offer a holistic look at the newsroom production cycle, review and compare journalistic mandates to those of information professionals, and identify crucial parameters in news verification with algorithmic solutions.

2. Literature Review

In this section, I review seemingly disparate sets of criteria for libraries and newsrooms which share commonalities in their efforts to provide accurate, complete, truthful, unbiased information. Each of these criteria should be seen as separate narrow-purpose tasks for automation with NLP and ML techniques. To provide an analogy, spell-checking was once a narrowly defined task for NLP research and it is now considered a practically solved problem.

2.1. Information Quality Criteria from Information Science Perspective

What does LIS have to offer for news verification in terms of criteria or parameters to strive for? Library science professionals curate information and provide access to credible sources. "Using library databases is a near-foolproof way to find credible information", the Harvard Library Research Guide introduces its databases ("Fake News, Misinformation, and Propaganda", 2017). Library mandates typically encompasses provision, filtering, selection, dissemination, and assistance in making sense of knowledge. Management Information Systems (MIS) assess information quality based on its usefulness or "fitness for use" (e.g., Knight and Burn (2005); Lee, Strong, Kahn, and Wang (2002); Stvilia, Al-Faraj, and Yi (2009)). Various authors stress components in verification that could be defined as a narrow task for NLP: accuracy, believability, reputation, objectivity (Wang & Strong, 1996), factuality (Zmud, 1978), credibility, consistency and completeness (Jarke & Vassiliou, 1997), correctness and lack of ambiguity (Wand & Wang, 1996), and precision, reliability, freedom from bias (DeLone & McLean, 1992).

"The [LIS] profession's deep commitment to verified sources and reliable information mirrors similar values – accountability for accuracy, careful research before drawing firm conclusions, and a willingness to correct errors – that drive responsible journalism." ("Fighting Fake News", 2016) Since the latest surge of "fake news" around 2016 US Presidential Elections, LIS professionals actively step-by-step digital literacy guidelines (see Figure 1) ("How to spot fake news", 2017).

Figure 1. IFLA's Infographic with 8 Steps to Discover the Verifiability of a Given News-Piece (based on FactCheck.org's "How to Spot Fake News", 2016).

"Like it or not, most people get the vast majority of their information about the world from the news", reporters emphasize ("Introducing Factmata", 2016). And LIS responds by reinforcing the role of digital literacy in the age of increased online misinformation:

> "Perhaps now more than ever, digital literacy matters, not just as a skill set to enable citizens to analyze and evaluate the information they encounter; it is critical to many important social phenomena, including positive health outcomes…, workforce development, and participative governance." (Webinar, 2017)

2.2. Journalistic Criteria of Excellence

Kovach and Rosenstiel (2010) instruct *journalists* to ask themselves the following six key questions regarding veracity and quality of information in their investigative process:

1. What kind of content am I encountering?
2. Is the information complete, and if not, what is missing?
3. Who and what are the sources and why should I believe them?
4. What evidence is presented, and how was it tested or vetted?
5. What might be an alternative explanation to understanding?
6. Am I learning what I need to? (Kovach & Rosenstiel, 2010, p. 32)

According to other classic journalistic textbooks, *editors* should assess news for newsworthiness with these criteria in mind: big picture composition, completeness, conflict, consequence, continuity, continuity, controversy, elite countries and persons, importance, interest, negativity, noteworthiness, novelty, oddity, predictability, relevance, sensationalism, timeliness, unexpectedness, veracity (Galtung & Ruge, 1973; Shoemaker & Mayfield, 1987).

Figure 2 describes the workflow of traditional journalists with an intermediate step of fact-checking and editorial staff who verifies facts, events, sources, which typically allows for production of higher quality news stories. It is then contrasted to the more immediate "news sharing a.s.a.p." model in which many newsroom professional functions disappear due to staff cuts (Kovach & Rosenstiel, 2010).

3. News Verification with NLP and ML

The idea is to implement a complex of discrete tasks, state-of-the-art in NLP and ML – *the News Verification Suite*, as informed by LIS information quality criteria and best newsroom practices. In Figure 2, I propose technological interventions for news content verification at the 3 stages of the workflow: sources → content producers, content producers → newsreaders, and in social media channels. This design offers discrete measurements along textual and networking parameters – such as novelty, negativity, veracity, presence of satire, bias, conflict, promotional material. Here is an operational illustration. A news reporter scans eyewitness accounts for falsehoods with Deception Detection, then uses semantic analysis to extract previously related events, and assess the novelty of the news with Topicality and Novelty Detection. An editor employs Sentiment Analysis to fine-tune the style and rhetoric of the reporter and screens out Sensationalist statements. Information curators search for multiple news stories on the same topic and Auto-Summarize multiple perspectives, scanning for fraudulent information and suspicious sources. Newsreaders may apply similar criteria to fully engage in critical thinking. Alerts are produced with Rumor Debunking, Satire and Native Ads Labeling.

The overall effect of technological intervention is increased awareness of potential pitfalls in the digital news sharing process. These steps should shadow the decision-making processes of the reporters, editors, fact-checkers which are generally thought of as unattainable with artificial intelligence as of now. Once a set of feasible NLP sub-tasks are clearly defined, they are bound to improve with time and R&D.

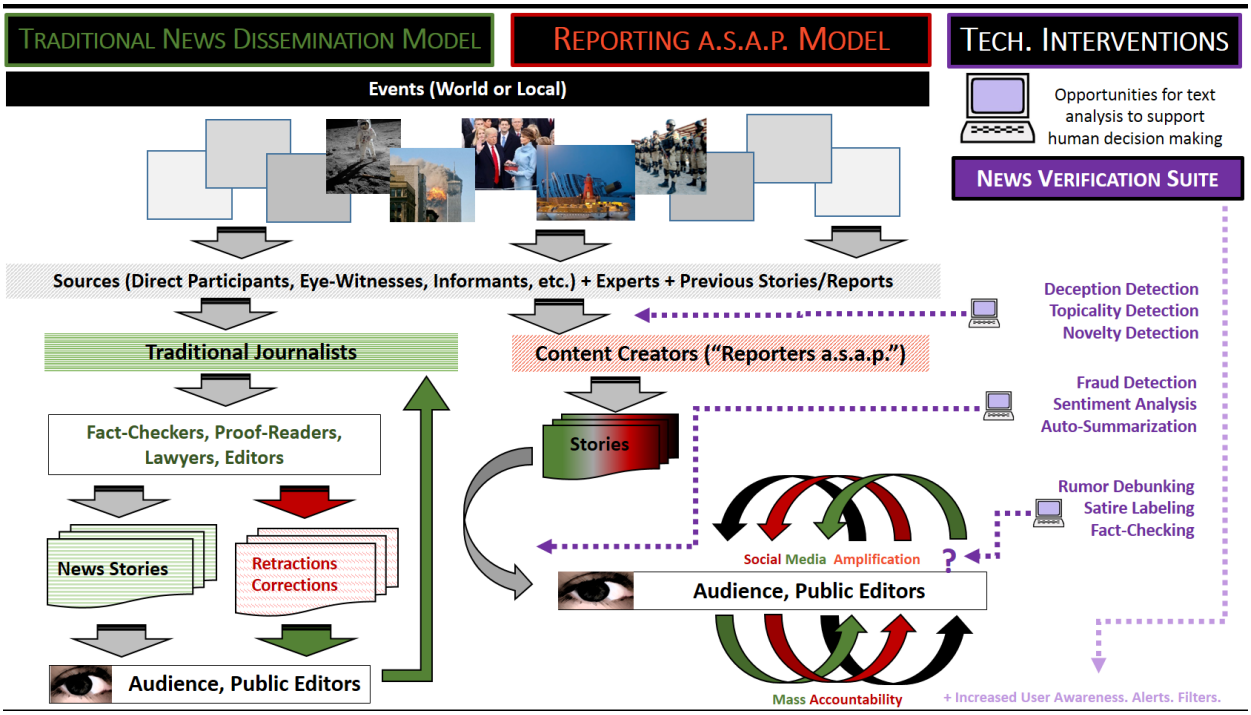

Figure 2. Model of Traditional vs. Digital A.S.A.P. Reporting Workflow. Potential areas for intervention with NLP technologies are indicated. The goal is to support human discernment in news consumption, digital content curation and decision-making.

4. Conclusions

As of 2017, varieties of "fake news" proliferate in digital environments but each type requires its own solution. I propose a conceptual vision for a multi-layers NLP system design, *the News Verification Suite*, which combines concrete technological interventions in the digital news cycle, in correspondence with excellence criteria in newsroom reporting and LIS information quality assessments. I call for development of automated solutions to previously undefined AI problems. This research agenda aims to identify and implement a set of feasible NLP tasks to support the work of reporters, editors, and newsreaders in verifying suspicious instances of news and in broader fact-checking. Further educational measures for newsreaders are undoubtedly necessary to heighten information literacy and critical thinking skills.

5. 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. Rubin for the project entitled *Digital Deception Detection: Identifying Deliberate Misinformation in Online News.* This work could not have been possible without the dedicated team has been contributing towards a working system based on this vision. Special thanks for stimulating conversations to Paul Benedetti, an investigative journalist and a faculty member at Western, and my graduate research assistant team: Niall Conroy, Yimin Chen, Sarah Cornwell, and Toluwase Asubiaro.

6. References

Chen, Y., Conroy, N. J., & Rubin, V. L. (2015a). *Misleading online content: Recognizing clickbait as "false news".* Paper presented at the Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection.

Chen, Y., Conroy, N. J., & Rubin, V. L. (2015b). *News in an online world: The need for an automatic crap detector.* Paper presented at the Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community.

Cornwell, S. E. L., & Rubin, V. L. (2017). *Native advertising: Ads in disguise as editorials*. Paper presented at the 45th Annual Conference of the Canadian Association of Information Science, Toronto, Ontario, Canada.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information systems research, 3*(1), 60-95. doi:10.1287/isre.3.1.60

Donald trump's alarming alternative facts. (2017, January 23, 2017). Editorial. *The Boston Globe*. Retrieved from https://www.bostonglobe.com/opinion/editorials/2017/01/23/ donald-trump-alarming-alternative-facts/sbmPy9m5wx4pMEzp4tdNnK/story.html

Fake News, Misinformation, and Propaganda. (2017). *Harvard Library Research Guide*. Retrieved from http://guides.library.harvard.edu/fake.

Fox, C. J. (1983). Information and misinformation: An investigation of the notions of information, misinformation, informing, and misinforming. Westport, CT: Greenwood Press

Galtung, J., & Ruge, M. (1973). Structuring and selecting news. *The manufacture of news: Social problems, deviance and the mass media, 1*(62), 62-72.

Hancock, J. T. (2012). Digital deception: When, where and how people lie online. In A. N. Joinson, K. McKenna, T. Postmes, & U.-D. Reips (Eds.), *The oxford handbook of internet psychology* (pp. 508): Oxford University Press.

Fighting fake news: How libraries can lead the way on media literacy (2016, December 27, 2016). *American Libraries Magazine*. Retrieved from https://americanlibrariesmagazine.org/2016/12/27/fighting-fake-news/

How to spot fake news (2017, 18 April 2017). *IFLA Publications.* Retrieved from https://www.ifla.org/publications/node/11174

How to spot fake news (2016, 18 November, 2016). *FactCheck.org Website.* Retrieved from http://www.factcheck.org/2016/11/how-to-spot-fake-news/

Introducing Factmata: Artificial intelligence for automated fact-checking (2016, 6 December, 2016). *Factmata Website*. Retrieved from: https://medium.com/factmata/introducing-factmata-artificial-intelligence-for-political-fact-checking-db8acdbf4cf1

Jarke, M., & Vassiliou, Y. (1997). *Data warehouse quality: A review of the dwq project.* Paper presented at the Conference on Information Quality, Massachusetts Institute of Technology, Cambridge MA USA.

Knight, S.-a., & Burn, J. M. (2005). Developing a framework for assessing information quality on the world wide web. *Informing Science Journal, 8*(5), 159-172.

Kovach, B., & Rosenstiel, T. (2010). *Blur: How to know what's true in the age of information overload* (Vol. 1st U.S.). New York: Bloomsbury.

Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). Aimq: A methodology for information quality assessment. *Information & Management, 40*(2), 133-146.

Rieh, S. Y. (2010). Credibility and cognitive authority of information. In M. J. Bates (Ed.), *Encyclopedia of library and information sciences* (Vol. 2, pp. 1337-1344): Taylor & Francis.

Webinar: The Role of Information Scholars and Professionals in Responding to Fake News, Misinformation, and Propaganda. (2017, 31 March 2017). *ASIST Events.* Retrieved from: https://www.asist.org/events/webinar-the-role-of-information-scholars-and-professionals-in-responding-to-fake-news-misinformation-and-propaganda/

Rubin, V. L. (2010). *On deception and deception detection: Content analysis of computer-mediated stated beliefs.* Paper presented at the 73rd ASIS&T Annual Meeting: Navigating Streams in an Information Ecosystem, Pittsburgh, Pennsylvania.

Rubin, V. L. (2017). Deception Detection and Rumor Debunking for Social Media. In Sloan, L. & Quan-Haase, A. (Eds.) (2017) *The SAGE Handbook of Social Media Research Methods*, London: SAGE. https://uk.sagepub.com/en-gb/eur/the-sage-handbook-of-social-media-research-methods/book245370

Rubin, V. L., Chen, Y., & Conroy, N. J. (2015). *Deception detection for news: Three types of fakes.* Paper presented at the Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community.

Rubin, V. L., Conroy, N. J., & Chen, Y. (2015, January 5-8, 2015). *Towards news verification: Deception detection methods for news discourse.* Paper presented at the Hawaii International Conference on System Sciences (HICSS48) Symposium on Rapid Screening Technologies, Grand Hyatt, Kauai, Hawai'i.

Rubin, V. L., Conroy, N. J., Chen, Y., & Cornwell, S. (2016, June 17, 2016). *Fake news or truth? Using satirical cues to detect potentially misleading news.* Paper presented at the The 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, CA, USA.

Shoemaker, P. J., & Mayfield, E. K. (1987). Building a theory of news content: A synthesis of current approaches. *Journalism and Communication Monographs, 103*.

Stvilia, B., Al-Faraj, A., & Yi, Y. J. (2009). Issues of cross-contextual information quality evaluation—the case of arabic, english, and korean wikipedias. *Library & Information Science Research, 31*(4), 232-239. doi:10.1016/j.lisr.2009.07.005

Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM, 39*(11), 86-95. doi:10.1145/240455.240479

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5-33.

Zmud, R. W. (1978). An empirical investigation of the dimensionality of the concept of information. *Decision Sciences, 9*(2), 187-195. doi:10.1111/j.1540-5915.1978.tb01378.x