The Spread and Mutation of Science Misinformation

Ania Korsunska(✉)

Syracuse University, Syracuse, USA
akorsuns@syr.edu

Abstract. As the media environment has shifted towards digitization, we have seen the roles of creating, curating and correcting information shift from professional "gatekeeper" journalists to a broader media industry and the general public. This shift has led to the spread of misinformation. Though political "fake news" is currently a popular area of study, this study investigates another related phenomenon: science misinformation. Consistent exposure to science misinformation has been shown to cultivate false beliefs about risks, causes and prevalence of illnesses and disincentivize the public from implementing healthy lifestyle changes. Despite the need for more research, science misinformation dissemination studies are scarce. Through a case study that traces the spread of information about one specific article through hyperlink citations, this study adds valuable insights into the inner workings of media networks, conceptualizations of misinformation spread and methodological approaches to multi-platform misinformation tracing. The case study illustrates the over-reliance of media sources on secondary information and the novel phenomenon of constantly mutating online content. The original misinformant is able to remove misleading information, and as a result, all of the subsequent articles end up referencing misinformation to a source that no longer exists. This ability to update content online breaks the information flow process: news stories no longer represent a snapshot in time but instead living, mutating organisms, making any study of media networks increasingly complex.

Keywords: Science misinformation · Network analysis · Network gatekeeping

1 Introduction

Traditionally, journalists in the U.S. have been called upon to inform the public of the new and important issues of the day. They are content producers, as well as gatekeepers of information, deciding what is worthy of coverage (Lewin 1947; Shoemaker and Vos 2009). As the media environment has shifted towards digitization, we have seen the roles of creating and curating information, and correcting misinformation spread throughout a broader media industry as well as the public (Barzilai-Nahon 2008; Lazer et al. 2018). This shift in the media environment has helped give rise to the creation and dissemination of misinformation through oversimplification and exaggeration (Woloshin and Schwartz 2002; Lewandowsky et al. 2012; Howell 2013). Though political "fake news" is currently a popular area of study (Allcott and Gentzkow 2017;

N. G. Taylor et al. (Eds.): iConference 2019, LNCS 11420, pp. 162–169, 2019.
https://doi.org/10.1007/978-3-030-15742-5_15

Flynn et al. 2017; Vosoughi et al. 2018), this study aims to draw attention to another related phenomenon: science misinformation.

2 The Impact of Misinformation

Misinformation, defined as the unintentional spread of false, inaccurate or misleading information (Kumar and Geethakumari 2013; Lazer et al. 2018) about scientific studies is abundant. It is false, because it is contradicted by the best available evidence in the public domain - the scientific consensus. Misinformation is virtually unavoidable in our current media environment (Moynihan et al. 2000; Schwitzer 2008, 2017) and can have lasting, negative effect on people's lives. Consistent exposure to misinformation related to science and health has been shown to cultivate false beliefs about risks, causes and prevalence of illnesses and disincentivize the public from implementing healthy lifestyle changes and getting routine checkups (Niederdeppe and Levy 2007). Misinformation has also been shown to spread much faster and wider than the truth (Vosoughi et al. 2018), and be very difficult to eliminate from people's minds, even after being debunked (Hochschild and Einstein 2015; Nyhan and Reifler 2015). Despite the dire need, multi-platform approaches to the study of science misinformation dissemination are few and far between. Though current research has looked at the spread of misinformation within platforms where the flow can be easily traced, such as sharing/retweets on Twitter or Facebook (Vosoughi et al. 2018; Allcott and Gentzkow 2017; Wu and Liu 2018; Vraga and Bode 2018), to my knowledge there has not been an attempt to develop a method of tracing information in relation to specific topics across the Internet.

This case study proposes the idea of hyperlink citation tracing, which enables us to see the flow of information in between various media sources: peer-review journals, press release websites, traditional media and "new" media (blogs, social media, etc.). With this method we will be able to understand the information's life cycle: creation, spread, mutation, and possibly death. In this study, I ask the following questions: (1) How does misinformation flow in between different media sources? (2) Which media outlets rely on primary sources and which on secondary sources? (3) Which media outlet is the source of the misinformation? Answering these questions in the context of this case study will helps us understand the complexity of the misinformation spread and is the first step to developing methodological approaches to multi-platform misinformation tracing, as well as future misinformation prevention algorithm development.

3 Scientists Say Smelling Farts Might Prevent Cancer

Published in the *MedChemComm Journal* on April 7, 2014, the article at the heart of this case study is titled "The synthesis and functional evaluation of a mitochondria-targeted hydrogen sulfide donor(10-oxo-10-(4-(3-thioxo-3H-1,2-dithiol-5-yl) phenoxy) decyl) triphenylphosphonium bromide (AP39)" (Le Trionnaire et al. 2014). On July 9, 2014 the press release was published by the University of Exeter, with which 6 out of 8

authors are affiliated. It is titled "Rotten egg gas holds key to healthcare therapies" (University of Exeter 2014). In its introduction the authors state that even though "it may smell of flatulence" hydrogen sulfide "is now being found to offer potential health benefits in a range of issues, from diabetes, to stroke, hearts attacks and dementia" (University of Exeter 2014). Though *the Daily Mail* was the first popular media publication to report on this study, it did not spread the information very far. *Time* published an article on July 11, 2014 with the unfortunate title "Scientists Say Smelling Farts Might Prevent Cancer" (Stampler 2014a). Three days later on July 14, the title was amended, removing the scientists: "Ridiculous study of the day says smelling farts might prevent cancer" (Stampler 2014b). On July 18 the was fully rewritten and retitled "A stinky compound may protect against cell damage, study finds" (Stampler 2014c). The author included a correction notice: "An earlier version of this article incorrectly summarized the findings and implications of this study" (Stampler 2014c). Once *Time* introduced specific elements to the story, such as cancer and farts, neither of which is mentioned in the original paper or the press release, the story spread rapidly.

4 Methods: Building the Citation Network

This case study analyzed the flow of information by utilizing hyperlink citations of popular media publications. The Google search engine was used as a data source, with keywords "farts," "cancer" and "study" and date range July 1–31, 2014. Though the study was published in April, there was no media coverage until the press release in July. The first six pages of Google results were scraped, excluding duplicates and stories not directly related to the topic (n = 48). Including the peer-review publication and the press release, total sample size was 50 articles. From each popular media article all external, working hyperlinks, article titles, URLs and dates of publication were extracted. Titles were coded as to whether they were proliferating the misinformation or taking an opposing, fact-checker role. The hyperlinks were converted into numerical values and coded as edges, with the transfer of information from source to target. The network was visualized with Gephi, an open-source network analysis software (Bastian et al. 2009). Once it became clear that *Time* produced multiple versions of the same article, additional analysis was conducted by utilizing Archive.org, the digital library of Internet content.

The resulting network contained 71 nodes (total sources cited) and 123 edges (total hyperlink citations). The most important nodes were the peer review publication *MedChemComm Journal* (cited by 13 sources), the University of Exeter press release (cited by 24 sources) and *Time* (cited by 22 sources). Out of the 49 articles (excluding the peer-review publication), 37 (76%) were proliferating the misinformation in their titles (Fig. 1).

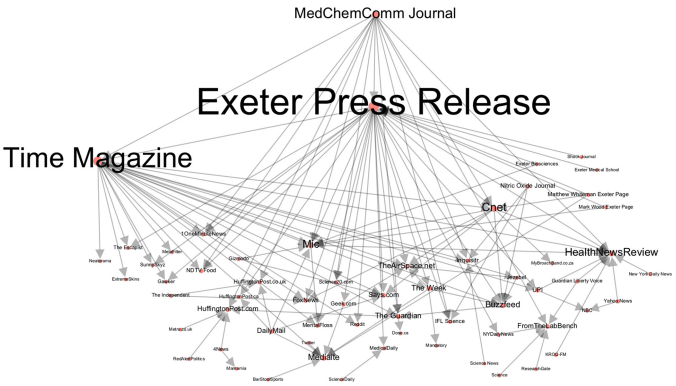

Fig. 1. Information spread network. Nodes represent article sources, edges – hyperlink citations, and arrows - the directionality of information transfer (from source to target)

5 Analysis: Tracing Misinformation Flow

First, to broadly look at the flow of misinformation in between the different media sources. Out of the sample, 12 articles were published on the same day as the *Time* article, all with titles regurgitating the misinformation; 10 were published the next day (with 7 misleading titles). From there, coverage dwindles to just 8, 8, 3 and 1 throughout the next four days. Curiously, 7 articles were published after *Time* rewrote their article on 7/18/2014, with 5 continuing to spread misinformation. Those 5 listed either no sources, no reliable sources (small, unknown blogs) nor did they reference *Time*. The 2 skeptical articles listed the *MedChemComm Journal,* among others, in their hyperlink citations. This is consistent with literature that shows that misinformation spreads much faster and wider than the truth (Vosoughi et al. 2018), making the job increasingly more difficult for both official fact-checking organizations such as FactCheck.org and Snopes, and for popular media publications that are trying to correct the story, to correct misinformation once it spreads throughout media networks.

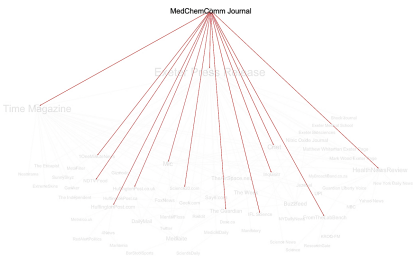

Fig. 2. *MedChemComm Journal* article spread

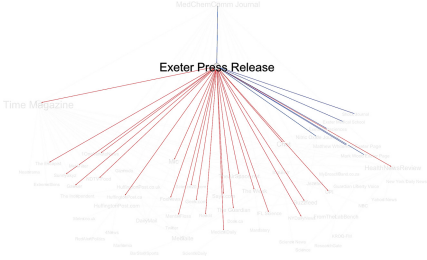

Fig. 3. University of Exeter press release spread

It is evident that popular media over-rely on secondary sources, such as The University of Exeter press release, *Time*, etc., as opposed to the original scientific publication. With the help of visualizations, we see the difference in the spread of the information directly from the original *MedChemComm* publication, cited by 13 sources (Fig. 2), as compared to the much wider scope of spread of information of the University of Exeter press release, cited by 24 sources (Fig. 3). The popular media publications refer to each other as sources of information, creating a type of echo-chamber. With their hyperlink language they also illustrate a need to transfer responsibility and accountability of fact checking one's sources prior to publication. It can be argued that this practice stems from the need for transparency of sources as a professional norm in journalism (Humprecht and Esser 2016). But when contextualized in the fast-paced online media world, it is much more likely a short-cut utilized to avoid taking responsibility for publishing unverified information. This is likely due to the de-professionalization of science journalism and overuse of amateur, freelance writers (Schwitzer 2008) coupled with the need to publish high quantity with high speed, little emphasis on quality or depth (Usher 2014).

Another interesting and novel theme that comes to light through this case study is the live mutation of information. The culprit in this study, *Time*, is a trusted, traditional media gatekeeper. By looking at just *Time* citations data (Fig. 4), we see that half of the other publications referenced it as a source of information (49%). It can be inferred that the authors of publications were influenced by *Time*'s reputation and attributed credibility to them as a source. Individuals' evaluations of the strength of an argument largely rely on heuristics and their perception of source credibility (Hovland and Weiss 1951; Kahan et al. 2010). If traditional gatekeepers do not live up to their reputation, they stand to lose the public's perception of credibility.

Unfortunately, in this case, *Time* was the source of misinformation. By looking at Archive.org, we can see that *Time* published misinformation and then was able to remove it completely. This is a novel complication to the media spread network structure, as all of the references to the *Time* were made prior to the article being completely rewritten on 7/18/2014. The hyperlink citations as they stand now in all the blog posts are actually referencing information that no longer exists.

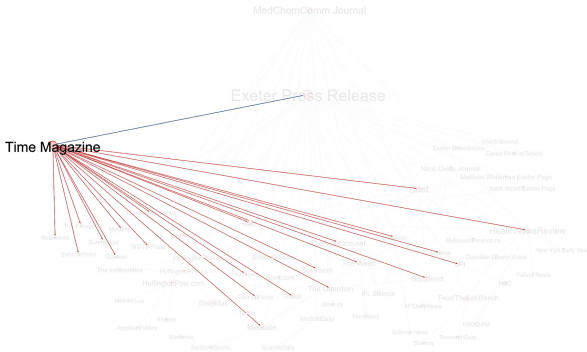

Fig. 4. *Time Magazine* article spread

This is a finding with significant implications for any further research in this subfield: the ability to update content online breaks the information flow process. News stories no longer represent a snapshot in time but instead living, mutating organisms (Fass and Main 2014).

6 Conclusion

The recent rise of media technology has meant an unprecedented increase in the speed and scope of the public's access of information, impacting how information flows in the media, as well as who creates and curates it (Hesse et al. 2005; Cline and Haynes 2001; Panth and Acharya 2015). There is more information available than ever before, with varying degrees of accuracy. Articles published online have a special opportunity: to not only reference their sources, but to directly hyperlink the sources to the original content. This is a style of citing and referencing that not only accomplishes the traditional goals of academic citations, but also allows us to see the flow of information in this unique, fast-paced media environment. This becomes especially important in the current age of "fake news" as we move towards what some have referred to as a "post-truth" world (Gewin 2017).

The aim of this case study was to present a sample approach to studying misinformation: the idea of investigatory hyperlink tracing, which enables us to see the entire flow of information through various media sources, and to draw attention to the issue of science misinformation and the real-time mutation of information online. By its nature as a case study with a limited sample, I make no claims as to generalizability of these themes. Utilizing the Google search was not ideal, as the "top" search results may be adjusted to the user and change according to the popularity of the websites at the time of search. Despite these limitations, the Google sample offered the closest available snapshot of what search results were available to general public at the time. Furthermore, coding was done without utilizing any computational software, so the sample size and analysis was limited. This case study, though, provides a myriad of future research ideas. Understanding the complexities of the misinformation spread process and various influencing factors involved can greatly aid in any larger-scale analysis utilizing machine learning algorithms, which is very timely in our current media environment. Additional research is needed to compare hyperlink tracing to other network analysis methods in order to see how useful this approach will be in a large-sale study. Further research should also focus on the evaluation of general trends of information spread throughout the media and continue to trace the life cycle of science misinformation back to its sources. This is the first step in developing methods for the detection or prevention of science misinformation spread, one of the most salient problems of our time.

References

Allcott, H., Gentzkow, M.: Social media and fake news in the 2016 election. J. Econ. Perspect. **31**(2), 211–236 (2017)

Barzilai-Nahon, K.: Toward a theory of network gatekeeping: a framework for exploring information control. J. Assoc. Inf. Sci. Technol. **59**(9), 1493–1512 (2008)

Bastian, M., Heymann, S., Jacomy, M.: Gephi: an open source software for exploring and manipulating networks. In: ICWSM, vol. 8, pp. 361–362 (2009)

Cline, R.J., Haynes, K.M.: Consumer health information seeking on the Internet: the state of the art. Health Educ. Res. **16**(6), 671–692 (2001)

Fass, J., Main, A.: Revealing the news: how online news changes without you noticing. Digit. Journal. **2**(3), 366–382 (2014)

Flynn, D.J., Nyhan, B., Reifler, J.: The nature and origins of misperceptions: understanding false and unsupported beliefs about politics. Polit. Psychol. **38**(S1), 127–150 (2017)

Gewin, V.: Communication: post-truth predicaments. Nature **541**(7637), 425–427 (2017)

Hesse, B.W., et al.: The impact of the internet and its implications for health care providers: findings from the first health information national trends survey. Arch. Intern. Med. **165**, 2618–2624 (2005)

Hochschild, J.L., Einstein, K.L.: Do Facts matter? Information and misinformation in American politics. Polit. Sci. Q. **130**(4), 585–624 (2015)

Hovland, C.I., Weiss, W.: The influence of source credibility on communication effectiveness. Public Opin. Q. **15**(4), 635–650 (1951)

Howell, L.: Digital wildfires in a hyperconnected world. WEF Report (2013). http://reports.weforum.org/global-risks-2013/risk-case-1/digital-wildfires-in-a-hyperconnected-world/. Accessed 14 Aug 2018

Humprecht, E., Esser, F.: Mapping digital journalism: comparing 48 news websites from six countries. Journalism **19**(4), 500–518 (2016)

Kahan, D.M., Braman, D., Cohen, G.L., Gastil, J., Slovic, P.: Who fears the HPV vaccine, who doesn't, and why? an experimental study of the mechanisms of cultural cognition. Law Hum. Behav. **34**(6), 501–516 (2010)

Kumar, K.K., Geethakumari, G.: Information diffusion model for spread of misinformation in online social networks. In: 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1172–1177. IEEE, August 2013

Lazer, D.M.J., et al.: The science of fake news. Science **359**(6380), 1094–1096 (2018)

Le Trionnaire, S., et al.: The synthesis and functional evaluation of a mitochondria-targeted hydrogen sulfide donor, (10-oxo-10-(4-(3-thioxo-3 H-1, 2-dithiol-5-yl)phenoxy)decyl) triphenylphosphonium bromide (AP39). MedChemComm **5**(6), 728–736 (2014)

Lewandowsky, S., Ecker, U.K., Seifert, C.M., Schwarz, N., Cook, J.: Misinformation and its correction: continued influence and successful debiasing. Psychol. Sci. Public Interest **13**(3), 106–131 (2012)

Lewin, K.: Frontiers in group dynamics: II. Channels of group life, social planning, and action research. Hum. Relat. **1**(2), 143–153 (1947)

Moynihan, R., et al.: Coverage by the news media of the benefits and risks of medications. N. Engl. J. Med. **342**(22), 1645–1650 (2000)

Niederdeppe, J., Levy, A.G.: Fatalistic beliefs about cancer prevention and three prevention behaviors. Cancer Epidemiol. Prev. Biomark. **16**(5), 998–1003 (2007)

Nyhan, B., Reifler, J.: Displacing misinformation about events: an experimental test of causal corrections. J. Exp. Polit. Sci. **2**(1), 81–93 (2015)

Panth, M., Acharya, A.S.: The unprecedented role of computers in improvement and transformation of public health: an emerging priority. Indian J. Commun. Med. Official Publ. Indian Assoc. Prev. Soc. Med. **40**(1), 8 (2015)

Schwitzer, G.: How do US journalists cover treatments, tests, products, and procedures? an evaluation of 500 stories. PLoS Med. **5**(5), e95 (2008). 0700–0704

Schwitzer, G.: Pollution of health news. BMJ **1262**, j1262 (2017)

Shoemaker, P.J., Vos, T.P.: Gatekeeping Theory. Routledge, New York (2009)

Stampler, L.: Scientists Say Smelling Farts Might Prevent Cancer. Time Magazine. Internet Archive, WayBackMachine, 11 July 2014a. https://web.archive.org/web/20140711162609/http://time.com/2976464/scientists-say-smelling-farts-might-prevent-cancer/. Accessed 5 Aug 2018

Stampler, L.: Ridiculous Study of the Day Says Smelling Farts Might Prevent Cancer. Time Magazine. (2014b, July 14) Internet Archive, WayBackMachine. https://web.archive.org/web/20140714011647/http://time.com/2976464/scientists-say-smelling-farts-might-prevent-cancer/. Accessed 5 Aug 2018

Stampler, L.: A Stinky Compound May Protect Against Cell Damage, Study Finds. Time Magazine. Internet Archive, WayBackMachine, 18 July 2014c. https://web.archive.org/web/20140718153720/http://time.com/2976464/rotten-eggs-hydrogen-sulfide-mitochondria/. Accessed 5 Aug 2018

University of Exeter Research News: Rotten egg gas holds key to healthcare therapies, 9 July 2014. http://www.exeter.ac.uk/news/research/title_393168_en.html. Accessed 5 Aug 2018

Usher, N.: Making News at The New York Times. University of Michigan Press, Ann Arbor (2014)

Vosoughi, S., Roy, D., Aral, S.: The spread of true and false news online. Science **359**(6380), 1146–1151 (2018)

Vraga, E.K., Bode, L.: I do not believe you: how providing a source corrects health misperceptions across social media platforms. Inf. Commun. Soc. **21**(10), 1337–1353 (2018)

Woloshin, S., Schwartz, L.: Translating research into news. JAMA **287**(21), 2856–2858 (2002)

Wu, L., Liu, H.: Tracing fake-news footprints: characterizing social media messages by how they propagate. In: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, pp. 637–645. ACM (2018)