

# Public health, social networks, and the digital media ecosystem

## Emerging hypotheses

**Brittany Seymour**

*Harvard School of Dental Medicine*

*Berkman Klein Center for Internet and Society*

**Sands Alden Fish II**

*Center for Civic Media, MIT Media Lab*

**Fernando Bermejo**

*Center for Civic Media, MIT Media Lab*

**Natalie Gyenes**

*Berkman Klein Center for Internet and Society*

**Ethan Zuckerman**

*Center for Civic Media, MIT Media Lab*

**Hal Roberts**

*Berkman Klein Center for Internet and Society*

*The rapid far-reaching spread of misinformation has posed challenges for public health. Media Cloud, a global archive of more than 500 million online publications (mainstream media, blogs, advocacy pieces, original research articles, etc.) is an open source platform developed by the Harvard Berkman Klein Center for Internet and Society and MIT Center for Civic Media. Designed to evaluate the media's framing of issues online, the project maps information ecosystems publishing about given topics. To date, we have performed approximately 150 case studies using Media Cloud, including on public health topics such as vaccines and Ebola. Our work supports emerging hypotheses that online public health information ecosystems may behave more like complex social networks, and not merely unidirectional information distributors. Social norms are a strong determinant of group behavior, and future research can assist in addressing the determinants of online health misinformation by studying the dynamics of the digital media ecosystem.*

### **Keywords**

Misinformation, Health Communication, Internet, Digital Media, Social Networks

## Introduction: Digital Pandemics

Dr. Julio Frenk, while Dean of the Harvard Chan School of Public Health, proclaimed in a 2009 NIH Barmes lecture, “During the 20<sup>th</sup> century alone, the world as a whole experienced a larger gain in life expectancy than in all the previously accumulated history of humankind.” This increase is largely attributed to the success of vaccination, named one of the great public health achievements of the 20<sup>th</sup> century.<sup>1</sup> However, in 1998, a new publication alleged an association between the MMR vaccine and autism.<sup>2</sup> The authors disclosed in the article that they could not claim a causal link,<sup>2</sup> and the paper was eventually found to be faulty and was retracted.<sup>3</sup> Nonetheless, a digital assault on the safety and efficacy of vaccines was eventually triggered, one that began with autism and has since ballooned to numerous other concerns. Celebrities, bloggers, and advocacy groups continued to perpetuate myths and misinformation online.<sup>4</sup> Exemptions from the recommended vaccination schedule nearly doubled in California between 2007 and 2013,<sup>5</sup> and as a result, the U.S experienced the largest measles outbreak in nearly a generation, traced to a case in Disneyland.<sup>6</sup>

Digital assaults do not occur only for vaccines. A controversial article released in 2014 recommended that fluoride (another CDC ‘great’<sup>1</sup>) be designated as a developmental neurotoxin in children.<sup>7</sup> Numerous scientists and more recent studies have debunked the claim,<sup>8,9</sup> yet, communities continue to cite the article and related online misinformation as a reason to cease water fluoridation in their municipalities.<sup>10</sup> Misinformation about “quarantines” and “waiting periods” led a New Jersey school to request two healthy children who had moved from Rwanda, a nation with zero cases, stay home during the 2014 Ebola outbreak.<sup>11</sup> A story describing how pharmaceutical companies purposefully engineered the Zika virus to start the outbreak

for profit permeated online social networks.<sup>12</sup> These digital assaults, or perhaps more aptly named digital pandemics,<sup>13</sup> and their rapid far-reaching spread of misinformation have posed severe challenges for public health, and corrective information is consistently ineffective.

## Methods: The Media Cloud project

Media Cloud, a global archive of more than 500 million online publications (mainstream media, blogs, advocacy pieces, etc.) is an open source platform developed by the Harvard Berkman Klein Center for Internet and Society and MIT Center for Civic Media. The platform was developed to evaluate the media’s framing of issues online and to map the information sources publishing about given topics. To date, we have performed approximately 150 case studies using Media Cloud, on topics ranging from Islamophobia to substance abuse to climate change. For each case study, we developed a query using Boolean keyword strings within a defined date range. For example, in our vaccine case study, we searched for content online containing words that begin with the stem ‘vaccin’, from June 1, 2014 to March 1, 2015. We built on the initial set of content by identifying each hyperlink citation embedded within the text and downloading additional relevant content into our dataset. This hyperlink technique was repeated to produce a comprehensive ‘spidered’ database.

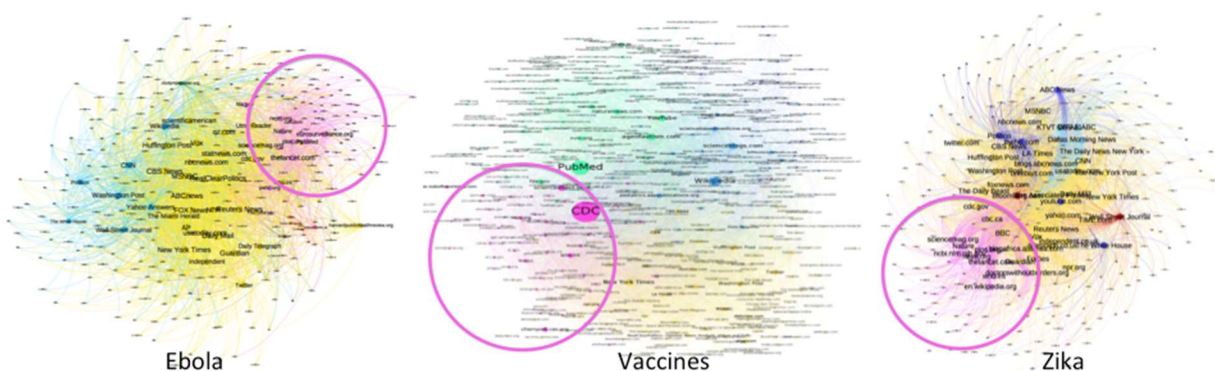
Media Cloud tools have been validated through previous studies and are now automated to allow for large volume data sets (i.e., several thousand to millions of publications). Research team members perform manual validation checks for each case study. Using this technology, we performed hyperlink and word frequency analyses to determine sources citing others through hyperlinks, the frequency with which

these citations happened, and which key terms were commonly used by sources in the information network generated. This allowed us to visually map communities of sources that both linked to one another most frequently, and shared common language to describe the topic at hand. Quantitative analysis of network characteristics allowed us to determine which sources and content were most influential within this media ecosystem for a given case topic.

## Results: Network clustering

We have detected preliminary patterns from our case studies. First, we identified sub-communities that formed within each topic

network, based on citations through hyperlinks between content publishers and common language (similar framing) used to describe the topic. Second, we detected information authorities — publishers that are cited or referenced most frequently by others in the network — within each sub-community. Third, we identified that information authorities in different sub-communities linked to one another less frequently than to sources within their own sub-communities. Fourth, each information network contained a distinct and isolated public health sub-community, primarily comprises public health authorities such as the CDC. Figure 1 illustrates three examples by topic of this digital network clustering.



**Figure 1.** Three example information networks for public health topics: Ebola, Vaccines, and Zika. Color-coded sub-communities within each topic network demonstrate which sources are linking to one another most frequently and sharing common language and framing of the topic. The distinct public health sub-communities within each information network are circled.

(Vaccine information network adapted from Getman R, Helmi M, Roberts H, Yansane A, Cutler D, Seymour B. Vaccine hesitancy and online information: The influence of digital networks. *Health Education and Behavior* 2017; Dec 1.)

## Discussion: Social networks in the digital media ecosystem

We were able to visualize these data in information network maps (Figure 1), where digital sources that link to one another most frequently, and utilize similar language to describe a given topic, cluster together. The tight linkage of these sub-communities may make it possible for existing opinions, based more on shared language and values than on fact, to propagate and reinforce each other. The public health sub-communities largely consisted of public health sources linking to other public health sources, which may mean, consequently, insufficient reach to audiences beyond the public health sub-communities. While public health sub-communities are helpful for building public health consensus, their closed nature may prevent the wider dissemination of evidence-based, and corrective information.

When investigating media ecosystems, we can identify information authorities deeply embedded within a given sub-community, indicating that their content is not reaching other sub-communities effectively. Information authorities likely generate a narrative around each topic that dominates their own sub-communities but not others. Our analyses have shown that information echoes from one source to another source within a given sub-community, even when counter to fact. Seemingly, information moves within the digital ecosystem based not on its 'scientific proof' but rather on its 'social proof,' a form of imitation where people ascribe to the behavior of their peers, such as sharing particular content, in order to resolve uncertainty.<sup>14</sup> These are perhaps the reflections of a shared set of social norms within sub-communities.

## Conclusion

Our case studies support emerging hypotheses that online public health information ecosystems may behave more like complex social networks, and not merely unidirectional information distributors. We are reminded of Geoffrey Rose's epidemiological thought piece, *Sick Individuals and Sick Populations*.<sup>15</sup> With sick individuals, he asks, "Why did *this* patient get *this* disease at *this* time." We look at the exposure, and the relative risk of contracting a disease compared to those who weren't exposed. Rose then asserts that for populations, we need to look not at individual risk, but at the determinants of incidence of disease for a population as a whole, and how incidences vary among different communities. "Why do some populations have high rates of a particular disease while in others, it is rare?" Initially, in the face of digital pandemics, we've been asking, "If the Internet is the exposure, what is the relative risk that someone will encounter misinformation?" However, perhaps we need to take Rose's advice and ask, "For a given topic, why might one online sub-community have higher rates of health misinformation than another? What distinguishes the two groups?" To find the determinants of the illness, he says we must study characteristics of populations, not individuals. He states social norms are a strong determinant of group behavior, and to shift the determinants of incidence of disease, in modern form, we must attempt to alter modern forms of society's behavioral norms.<sup>15</sup> Translating to online networks, we are hopeful the Media Cloud project can assist in addressing the determinants of misinformation by studying the dynamics of the digital media ecosystem.

## References

1. Centers for Disease Control and Prevention. 2012. Ten great public health achievements in the 20<sup>th</sup> century. Available at: <https://www.cdc.gov/mmwr/preview/mmwr.html/mm4850a1.htm>. Accessed Jan 5, 2018.
2. Wakefield, A.J., Murch SH, Anthony A, Linnell J, Casson DM, Malik M, et al. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children *Lancet* 351: 637-641, 1998.
3. The editors of the *Lancet*. Retraction--Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet*. 2010 Feb 6;375(9713):445.
4. Caulfield, Timothy, From Kim Kardashian to Dr. Oz: The Future Relevance of Popular Culture to Our Health and Health Policy (March 7, 2016). *Ottawa Law Review*, Vol. 47, No. 2, 2016. Available at SSRN: <https://ssrn.com/abstract=2744433> or <http://dx.doi.org/10.2139/ssrn.2744433>
5. Yang YT, Delamater PL, Leslie TF, Mello MM Sociodemographic Predictors of Vaccination Exemptions on the Basis of Personal Belief in California. *Am J Public Health*. 2016 Jan;106(1):172-7
6. Centers for Disease Control and Prevention. Measles — United States, January 1–May 23, 2014. *MMWR*. June 6, 2014 / 63(22):496-499.
7. Grandjean P, Landrigan PJ. Neurobehavioural effects of developmental toxicity. *Lancet Neurol*. 2014; 13(3):330-8.
8. Goldstein DA, Saltmires DA. Neurodevelopmental Toxicity: Still More Questions Than Answers. *The Lancet. Neurology*. England; 2014. P. 645.
9. Broadbent JM, Thomson WM, Ramrakha S, Moffitt TE, Zeng J, Foster Page LA, et al. Community Water Fluoridation And Intelligence: Prospective Study In New Zealand. *Am J Public Health*. 2015 Jan;105(1):72-6.
10. Mertz A, Allukian M. Community water fluoridation on the Internet and social media. *J Mass Dent Soc*. 2014;63(2):32-36.
11. Roberts H, Seymour B, Fish SA 2nd, Robinson E, Zuckerman E. Digital Health Communication and Global Public Influence: A Study of the Ebola Epidemic. *J Health Commun*. 2017;22(sup1):51-58.
12. Smith T. The Zika Conspiracies Have Begun. *Huffington Post*. February 5, 2016. Updated February 5, 2017. [http://www.huffingtonpost.com/tara-c-smith/the-zika-conspiracies-hav\\_b\\_9162854.html](http://www.huffingtonpost.com/tara-c-smith/the-zika-conspiracies-hav_b_9162854.html) Accessed April 10, 2017
13. Seymour B, Getman R, Saraf A, Zhang LH, Kalendarian E. When advocacy obscures accuracy online: digital pandemics of public health misinformation through an antifluoride case study. *Am J Public Health*. 2015;105(3):517-523. doi:10.2105/AJPH.2014.302437.
14. Berger J. *Contagious: Why Things Catch On*. New York, NY: Simon & Schuster; 2013.
15. Rose G. Sick individuals and sick populations. *Int J Epidemiol*. 2001 Jun;30(3):427-32; discussion 433-4.

---

## Contact the author

Brittany Seymour  
 Assistant Professor  
 Oral Health Policy and Epidemiology  
 Harvard School of Dental Medicine  
 188 Longwood Avenue, REB304  
 Boston, MA, USA 02115  
 Phone: +1 617-432-1848  
 Email: [brittany\\_seymour@hsdm.harvard](mailto:brittany_seymour@hsdm.harvard)