Draft Initial Analysis Part 1

Anti-vax and Bio-hacking Online Communities

Joshua Sinamo

1 Introduction

One of the marks in the beginning of the second millennium was the popularization of Web 2.0 (citation needed). This change has been allowing users to interact and share contents with fellow internet users. With the rise of user generated sites in Web 2.0 , more and more individuals seek health information from online news groups and blogs compared to the traditional evidence-based vaccine information sites (Witteman & Zikmund-Fisher, 2012; Kata, 2012; Chou et al., 2009; Lau et al., 2011). Intuitively, not all information available in the web as an important public source of health information is accurate; creating a concerning sphere for the spread of anti-vaccine sentiment (McDaid, Park (n.d.); Purcell et al., 2010; Fox & Duggan, 2013; Witteman & Zikmund-Fisher, 2012). Anti-vaccination activists able to inject their mes-

sages through the internet to the bigger audience, with a hope to gain new members (Hobson-West, 2007; Kitta, 2012).

Parents' fear of vaccines are not exclusively influenced by their individual beliefs. Smith in 2011 (as cited in Cooper, Larson & Katz (2008)) stated that parents who intentionally delayed vaccines for their child were significantly more likely to have heard or read negative information about vaccination when compared with parents who followed the recommended childhood vaccination schedule. Smith (2017) also expressed that the anti-vaccines "influencer" group has swayed the media in terms of vaccine perception, because the media are expected to serve balance reporting, which in return gave the group exposure to the public. Furthermore, nineteen percents of parents expressed concerns about vaccines, and the number jump to fifty percents in 2009 (Gellin & Maibach, 2000). We suspected that the flow of vaccine information online might shed some light to parents' acceptance to the idea of vaccination.

To explore more about the parents' hesitancy, Hagood & Herlihy (2013) categorized vaccine hesitancy into three type; vaccine rejectors, vaccine resistant, and vaccine hesitant. Other than complete rejection to the idea of vaccination, rejectors are more likely to utilize alternative medical practices and least likely to be open on to the education of vaccine awareness. Vaccine resistant on the other hand, are still willing to consider information about safety and benefit of vaccines (Wei et al., 2009; Healy, 2011). Vaccine-resistant parents' distrust of medical science is more rational than

vaccine-rejecting parents; there might be too many vaccines in the current schedule, or too soon for their child. Vaccine hesitant parents may vaccinate their children per health care provider' recommendation, but reluctant (Hagood & Herlihy, 2013). Many studies have also shown that even parents who vaccinate their children can have doubts and fears about immunization (Wilhelmina et. al., 2012; Hilton & Petticrew, Hunt, 2006; Kennedy et al., 2011; Cooper, Larson, Katz, 2008).

Regarding to the high proportions of vaccine hesitant parents, immunization experts are concerned that this population may shift from hesitancy to resistance, and resistance to unequivocal opposition (Schwartz & Caplan, 2011). Betsch et al. (2012) inferred that browsing an anti-vaccination website increased negative judgement towards immunization, but non-noticeable impact when viewing pro-vaccination ones. The result was further signified with the discovery that children of research subjects who viewed anti-vaccination sites had significantly lower vaccine coverage than its counterpart, five months after the study. In addition to that, Jones et al. (2012) also stated that "parents who used the Internet to get vaccination information were significantly less likely to consider healthcare providers and health authorities as trusted sources of vaccination information."

Among individuals who delayed or refused vaccines, they are significantly more likely to have looked for information related to vaccines on the internet (Dube et al., 2012; Smith et al., 2011). Parents who seek for the information about risk of childhood immunization will find more websites that perpetuate

myths about vaccines than those who seek for the information about the benefits of vaccination (Ruiz & Bell, 2014).

Anti-vaccines is not the only health-related movement which has been gaining spotlight in the Internet; biohacking is also one of them. According to Google Trends, the term biohacking was first used online in late 2008 and now encompasses a wide range of topics and activities. Bennett et al. (2009) and Delfanti (2013) defined biohacking as amateur practices of biological experimentation using a variety of do-it-yourself (DIY) tools and techniques in a non-traditional manner. Non-surprisingly, many biohackers are students of life sciences, have academic degrees and even hold academic positions besides their involvement in the biohacking community. According the Woodrow Wilson Center survey, biohackers are more educated than the general population and 19 per cent of them have received a doctorate degree (Grushkin et al. 2013).

In anti-vaccine communities, we discussed that fears and uneasiness towards the advancement of medicine technologies are the drive of the group; Conversely, in bio-hacking, there seems to be desires to pursue biological fitness through alternative means. Biohacking community comprised of mostly hobbyists founding molecular biology labs and doing biological engineering outside academia or other institutionalized settings (citation needed). The emerging community of biohackers has attracted the attention of public and academia, illustrated by mainstream media in an exaggerative manner with spotlights in its hope, hype, and horror (Seyfried, Pei & Schmidt 2014).

This spirit of innovation however, is not without a risk. Without a proper methodology as used in academic and industrial labs, biohacking practices can produce in highly dangerous outcomes. In the case of DIYbio (one of the institution for DIY biologist), Bennett et al (2009) and Heavey (2013) argued that although dangerous outcomes seems improbable, it is not impossible.

Bihani, Hartman, Sobiegalla, Rosenberg (2015), assessed two biohacking communities which utilizes online social media to share knowledge, expand their network and meet their respective goals: DIYbio and Bulletproof Executive. These two groups have online presence such as websites, blogs, facebook and twitter accounts. Bihani et. al (2015) also found that DIYbio as a non-commercial biohacking online-community has no form of public outreach for out-group individuals to learn about it. Home biohackers look to scientific articles for guidance and frequently link to them online, and DIY biologists rely on the infrastructure built up by synthetic biology (Roosth, 2010; Wexler, 2017). According to Delgado & Callen (2016), do-it-yourself biologists' goal is more political; power redistribution and the fundamental change of how science is done. An important similarity property between biohacking and anti-vaccines movement is their participatory approaches. Lengwiler (2008) refer to participatory approaches, as the involvement of non-scientists, laypeople, or citizens in science and technology. Genuine exchanges of ideas and bilateral communication between experts/stakeholders and general citizen are arguably encourage in participatory approaches, though in most cases are debatably still top-down, in which the depth and phase of engage-

ment are dictated by the top actors (Powell & Colin, 2008). Biohacking and anti-vaccines community in contrast, may be thought to challenge the top actors / experts. The process for both can take places outside traditional settings, though their motivations are vastly different (e.g. prevention of the improbable and pursuit of the improbable). Thus, biohacking seems to be more concerned with the progression of intellectual freedom and creativity, whereas anti-vaccines in conservation of the way of life.

In the Web 2.0. era, it was arguably hard to assess the flow of information received by a subset of people, considering that most communication in the internet happens in one-way street between news provider and audience. As the world became more and more modern, information also become easier to obtain. We hypothesized that considering the large number of people find their health related information online, hyperlink analysis is the suitable way to analyze the flow of online information. Hyperlink analysis provides the mapping of online sphere through hyperlink(s) between one webpage to another webpage

1.1 Research Question

- How does the structure and centrality measures of anti-vaccine and biohacking community changes as the depth of link increases, do they follow the law of large numbers (popular sites remain central or is there any hidden central communities inside the ecosystem)?

- How do opposing actors (agree vs disagree to the idea) interact overtime in the anti-vaccine and biohacking web communities (does pages tend to cluster together or is there any indication of infiltration from anti-vaccine actors to the pro-vaccine ecosystem)?

2 Methods

2.1 Data Source

Dataset is obtained through a self-made data miner that is able to mine the hyperlink data by reading through the html component of the page and find the out-link(s) from that particular link to another link. The seeds are obtained through google.com which utilizes PageRank algorithm to return the results for a given query. Law (2016), describes PageRank as an algorithm used by Google search engine to evaluate the importance of any individual webpage through gauging the number of relevant useful links between the page in question and other sites. Our data miner is inspired from issuecrawler.com. The depth limitation of VOSON (voson.uberlink.com) and IssueCrawler (issuecrawler.com) inspired us to create our own tool which has "no" depth limitation. Depth is defined as the number of link(s) from a seed to a particular website. Links collected is in a form of out-links exclusively. Besides than the difficulty of obtaining in-links, we argued that the number of in-links available at a certain time is never , because there's a delay in search engine indexing, even in the most prominent one like Google.com.

Seed links will be obtained from top 100 google searches per months for x(will be changed) years dating back to January 20xx(will be changed). These seed links then will be searched using the crawler for x(will be changed) depths. We will put labels according to the webpage agreement level to the topic with -1 as do not support, 0 as hesitant, and 1 as supportive. These

categories will be placed to each links in our data set using sentiment analysis algorithm that will be explained later in the paper.

2.2 Potential Obstacle

Potential obstacle includes time consuming data collection process. For example, if each webpage contains 10 meaningful links (excludes all self-links, etc), then for crawling depth = 100a, we will need to crawl 10^{100} links; running time increase exponentially as the depth increases. We argued that by increasing the depth, it will diminish the effect of trivial popularity (need source), so that we are able to see the popular community inside the community beyond the google browsing link, because inevitably these seeds are given based on popularity (insert citation). One way to tackle this problem is to split the seed links using multiprocessing. Heroku, python-selenium, and Graphene DB will be utilized to maximize running time.

2.3 Analysis

For visual analysis, We will utilize graph algorithms that is available in network graphing software Gephi. Numerical statistical analysis will be executed using NetworkX and other common numerical analysis package from python. (to be added).

References

TBA with proper APA citation. More resources to be added