



Selected Papers of #AoIR2019:
The 20th Annual Conference of the
Association of Internet Researchers
Brisbane, Australia / 2-5 October 2019

MAKING THE UNSEEN VISIBLE: EXPLORING CROSSCUTTING SOCIAL MEDIA PUBLICS AND THEIR SOCIOPOLITICAL TRAITS

Jakob Bæk Kristensen
University of Canterbury, New Zealand

Introduction

The potential of increased connectivity between people has been a main topic of discussion in media and communications research since the popularization of the internet, and especially since the web 2.0 revolution and the rise of social media. However, it is not easy to make claims about the effects and emergent properties of such digitally enhanced connectivity. Two somewhat clashing traits of digital media are 1) the increased complexity with which we find and appropriate information and form opinions about the world (Edgerly, Thorson, Bighash & Hannah, 2016) and 2) the increased amount of digital traces we leave behind when engaging with online materials (Jensen, 2013). While the first undoubtedly makes media and communication practices more complex to comprehend, the second should theoretically provide additional opportunities for reaching a deeper understanding of said complexity. However, bigger data does not always provide more accurate knowledge and researchers still need to make choices that greatly affect the results (boyd & Crawford, 2012). This paper combines socio-demographic and political classifications of social media users based on surveys with network clustering techniques in order to preserve the full-scale complexity of digital interactions while eliciting patterns relating to concrete sociopolitical properties.

Motivation and Previous Research

For the better part of a decade scholars have been using the digital breadcrumbs of peoples' online behavior to study media use in various contexts. This paper draws on two kinds of previous studies, those that make classifications and those that find clusters. An example of the latter is Bruns, Moon and Münch (2017) who mapped a significant subsample of the Australian Twittersphere showing the emergence and development of how certain groups of users cluster together based on how much they interact with each other. In that study however, there is little additional information about the users in each cluster. In contrast, other

Suggested Citation (APA): Kristensen, J. (2019, October 2-5). *Making the Unseen Visible: Exploring Crosscutting Social Media Publics and Their Sociopolitical Traits*. Paper presented at AoIR 2019: The 20th Annual Conference of the Association of Internet Researchers. Brisbane, Australia: AoIR. Retrieved from <http://spir.aoir.org>.

studies have shown that accurate classifications regarding personality, voting intention and various other socio-demographic features can be made from digital trace data (Youyou, Kosinski and Stillwell, 2015). The motivation behind this paper is that, although the purpose behind the two studies mentioned above are very different, the two approaches can be combined in a way that potentially allows one to explore previously unseen patterns. Retaining a lot of the complexity by mapping all interactions and looking for clusters that arise naturally from these while at the same time making predictions about different sociopolitical traits of the users provides the opportunity for looking at social media publics in greater detail.

Methods and Materials

This paper combines both supervised and unsupervised techniques to explore overlapping patterns of crosscutting publics on Facebook in Denmark. The first step involves using a survey of 1697 Danes whose public Facebook activity is coupled with their personal, self-reported information. Using machine-learning models, the survey is used to make predictions about the voting intention, gender, age and geo-location of a random sample of 50.000 Facebook users between 2016 and 2018. The activity of these 50.000 users are spread across 20.000+ public Facebook pages relating to politics, media, popular culture, music, art, unions and local venues. All their activity is modeled as a network and a clustering algorithm is used to divide users into groups based on which pages they interact most with compared to all other users and pages. This method thus allows one to compare the socio-political backgrounds that define the users with their belonging to a certain (naturally occurring) group based on the sum of their actual interactions on public Facebook pages.

Results

Results show that voting intention overlaps most strongly with the clusters in the network, followed by gender and then geo-location, while age has little overlap. They also show that clusters relating to the core political left, right and alt. right make up only smaller groups in the network with the majority of users having other interests as their main identification. Left voting users are overrepresented in clusters relating to lifestyle, NGOs and animals, the political right is more significantly interested in sport and infotainment shows. The alt. right, especially when compared to their size, completely dominates two clusters, one is the tabloid media and the other is a cluster made up of discussion around the most hot button political issues such as immigration, religion, crime and the EU. Furthermore the core alt. right cluster stands out by using angry emojis five times more on average than any other group.

In contrast to studies that tend to highlight polarization and echo-chamber formation on social media (e.g. Schmidt et al., 2017; Del Valle & Bravo, 2018), this mapping can provide a more balanced narrative. While the results of this study show that there are politically isolated clusters with up to 97% (core left, core right or core alt. right) users with the same voting intention, there are also

large clusters around mainstream news with a politically equal distribution of users. The main findings show the potential for exploring concrete sociopolitical trends in social media publics while conserving a high degree of the complexity of the digital trace data. While the method does not guarantee to elicit patterns that could not have been found with another approach, the way that it conserves complexity might make it less likely for researchers to overlook some patterns while overestimating the significance of others.

References:

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5), 662-679.

Bruns, A., Moon, B., Münch, F., & Sadkowsky, T. (2017). The Australian Twittersphere in 2016: Mapping the follower/followee network. *Social Media+ Society*, 3(4), 2056305117748162.

Del Valle, M. E., & Bravo, R. B. (2018). Echo Chambers in Parliamentary Twitter Networks: The Catalan Case. *International Journal of Communication*, 12, 21.

Edgerly, S., Thorson, K., Bighash, L., & Hannah, M. (2016). Posting about politics: Media as resources for political expression on Facebook. *Journal of Information Technology & Politics*, 13(2), 108-125.

Jensen, K. B. (2013). How to do things with data: Meta-data, meta-media, and meta-communication. *First Monday*, 18(10).

Schmidt, A. L., Zollo, F., Del Vicario, M., Bessi, A., Scala, A., Caldarelli, G. & Quattrociocchi, W. (2017). Anatomy of news consumption on Facebook. *Proceedings of the National Academy of Sciences*, 114(12), 3035-3039.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, 112(4), 1036-1040.