

# Community Detection in Google Searches Related to “Coronavirus”

PHILIP D. WAGGONER<sup>1,2,\*</sup>

<sup>1</sup>*Computational Social Science, University of Chicago, Chicago, IL, USA*

<sup>2</sup>*Institute for Social and Economic Research and Policy, Columbia University, New York, NY, USA*

## Abstract

Coronavirus and the COVID-19 pandemic have substantially altered the ways in which people learn, interact, and discover information. In the absence of everyday in-person interaction, how do people self-educate while living in isolation during such times? More specifically, do communities emerge in Google search trends related to coronavirus? Using a suite of network and community detection algorithms, we scrape and mine all Google search trends in America related to an initial search for “coronavirus,” starting with the first Google search on the term (January 16, 2020) to recently (August 11, 2020). Results indicate a near-constant shift in the structure of how people educate themselves on coronavirus. Queries in the earliest days focusing on “Wuhan” and “China”, then shift to “stimulus checks” at the height of the virus in the U.S., and finally shift to queries related to local surges of new cases in later days. A few communities emerge surrounding terms more overtly related to coronavirus (e.g., “cases”, “symptoms”, etc.). Yet, given the shift in related Google queries and the broader information environment, clear community structure for the full search space does not emerge.

**Keywords** *computational social science; data science; network science; text mining*

## 1 Introduction

The impact of COVID-19 on the global society has been striking. From isolation and fear of contracting coronavirus to the slowing of many local economies and limitation of in-person interactions, the effects of COVID-19 have threatened the modern flow of life. People across the world are often living and working in isolation, or at least are more separated from local communities than prior to the pandemic. Thus, how do people process such an unprecedented global crisis? In other words, absent in-person interaction, how do people self-educate and learn about coronavirus?

Despite the still-newness of COVID-19 and coronavirus, there is an impressive amount of extant work on related topics. Much of the research surrounding online behavior and COVID-19 tends to focus on misinformation (Pennycook et al., 2020; Bastani and Bahrami, 2020), fake news (Apuke and Omar, 2020; van der Linden et al., 2020), and smart-phone-enabled self-diagnosis (Collado-Borrell et al., 2020). Further, research on COVID-19 and online behavior is becoming increasingly creative and nuanced as the virus wears on, such as exploring online shopping (Laato et al., 2020), binge television watching (Dixit et al., 2020), and changes in sexual behavior (Lehmiller et al., 2020).

Yet, beyond explicitly pernicious or unique contexts, online knowledge dissemination (Chan et al., 2020) and information diffusion (Croce et al., 2020; Dinh and Parulian, 2020) surrounding

---

\* Email: [pdwaggoner@uchicago.edu](mailto:pdwaggoner@uchicago.edu).

the virus has also informed a good deal of recent work, such as followers and likes on social media (Ma et al., 2020), real-time scientific information distribution (Song and Karako, 2020), and even country-specific effects (Meier et al., 2020; Prem et al., 2020; Tuite Bogoch et al., 2020).

This body of rapidly developing research has also given rise to exploration of the social and behavioral effects of COVID-19, which are frequently evolving. For example, Kim and Bostwick (2020) explored the dimension of racial inequality, whereas Sher (2020); Gunnell et al. (2020) explored suicide and Venkatesh and Edirappuli (2020); Saltzman et al. (2020) explored mental health, all of which point to deepening and lasting effects of COVID-19 that will stick around long after the immediate threat of the virus is over. And still further, Waggoner (2020) placed policymaking on COVID-19 into historical context by focusing on the American context, which adds to a growing body of work on COVID-19 in a policymaking context (Cairney and Wellstead, 2020; Hartley and Jarvis, 2020).

In sum, whether focusing on news media (Liu et al., 2020), education and social networks (Elmer et al., 2020), infection and mortality rates (Baud et al., 2020), or forecasting (Anastasopoulou et al., 2020; Perc et al., 2020), the COVID-19 pandemic has captured the attention of the global research community. The result is an ever-growing store of research aimed at unpacking and understanding the far-reaching effects of the virus on the global population (Grech, 2020), policymakers (Waggoner, 2020), and medical workers (Grasselli et al., 2020).

Building on, though departing from some of the existing research on COVID-19 and online communities and building on recent similar work (Effenberger et al., 2020; Rovetta and Bhagavathula, 2020), we address this question by focusing on the individual level, starting at the place most people start when interested to learn more about any topic: Google. To construct an exploratory research design to address this question, we leverage a suite of network and community detection algorithms to mine Google search trends related to an initial search for “coronavirus.” We are interested in exploring that which people pair with the term “coronavirus” to understand how these individuals self-educate via Google (e.g., “coronavirus symptoms”, where “symptoms” is the term of interest).

Results across several stages show that Google behavior related to “coronavirus” have trended with various waves of the virus since its initial outbreak in December 2019. Further, there is clear structure around commonly associated terms such as “cases” and specific regions in the United States, such as “county.” Further, supporting results at a monthly interval of time demonstrate that while this clear aggregate structure exists at a relatively intuitive level, the structure of the Google search space has continued to evolve as time spent living with the virus has worn on. For instance, in the earliest days of the virus, search tended to focus on approaches to combatting coronavirus, whereas at the virus’s initial peak in March of 2020 in America, a spike in Googling of “stimulus checks” occurred. Community detection results show evidence of a few consistent communities in the aggregate, yet a rapidly evolving set of communities at the periphery, where sparse connections are uncovered.

## 2 Empirical Strategy

This analysis is focused on uncovering the structure of this Google search space of terms paired with an initial search for “coronavirus.” To take a step in understanding this structure, we first use text mining techniques to build a corpus of all scraped Google search queries in America related to an initial search for “coronavirus” from January 16, 2020 (coronavirus’s first Google query in the United States) to August 11, 2020.

## 2.1 Data: The Google Search Space

Prior to fitting and diagnosing networks of the Google search space related to “coronavirus,” it is useful to first explore the trend of interest in “coronavirus” over time to both validate selection of the primary search term (“coronavirus” over either “COVID” or “COVID-19”), while also visualizing the contours of relative interest in this term over time. To the latter point, the value is to place the Coronavirus into the context of the American search space of Google. In other words, visualization of Google searches for “coronavirus” relative to “COVID” or “COVID-19” shown in Figure 1, offers a baseline to move forward, and thus justifies focusing only on “coronavirus” moving forward in the remainder of the analysis, with the ability to generalize only for Google usage in America.

Figure 1 shows relative, or scaled interest in “Coronavirus” in the aggregate by day of highest level of interest (i.e., divide all days by most frequent day of Googling). A few trends are notable. First, as expected, there are no searches prior to December 2019, which makes sense given the reported beginning and spread of the virus in the global community beginning in December 2019. Also, Figure 1 shows the peak of interest in Googling the virus in America was late March/early February, with a drop in interest after this point. This point is when Coronavirus was at its initial height in America.

Building on the descriptive temporal trends in Figure 1, another helpful way to understand general patterns of Googling interest in “coronavirus” is to examine the geographic spread of the data. These results across the full study period are shown in Figure 2.

Surprisingly, in Figure 2, the highest relative interest in the United States are not in the (initial) hot-spot regions such as New York or Texas. Rather, some of the highest interest in “coronavirus” averaged across the full study period are in Idaho, New Mexico, and Michigan. This suggests that contraction of the virus is not necessarily linked to the interest in the virus. Though this is not a causal claim, these descriptive patterns in Figure 2 suggest that explicit probing of the connection between *interest* versus *virus contraction* may lead to unexpected, albeit interesting patterns. For present purposes, the relative interest in “coronavirus” appears to be geographically diffuse. Such an exploratory look at the Google search space, combined with the descriptive temporal pattern shown in Figure 1 sets the stage for diving into the Google searches explicitly to deepen an understanding of the structure of self-education via Google in the time of COVID-19.

## 2.2 Methods: Text Mining and Networks

In this research we are interested in mining these text data using a network science framework. We opt for this approach to explore the space for several reasons. The global use of and access to Google implies some meta-level of connection and community across the full user base. Related, As the users of Google are connected in general through a common interface, the use of Google to explore and learn about COVID-19 and coronavirus should also be community-based, where common threads in searching behavior should emerge given the common use of Google to self-educate on a variety of topics (e.g., Atlas et al., 2018; Aljilani and Kadobayashi, 2015; Ward et al., 2018). Also, community detection is a power approach to explore (i.e., with a lack of a clear causal question or framework as in our case) some space and learn from it, as with unsupervised machine learning, there are no sets of rules that define the searching and learning process in a formal way (Fortunato and Hric, 2016). As a result, the use of networks and community detection-based mining to home in on commonalities and differences across the Google search space is a reasonable approach to explore and learn the structure of some space.

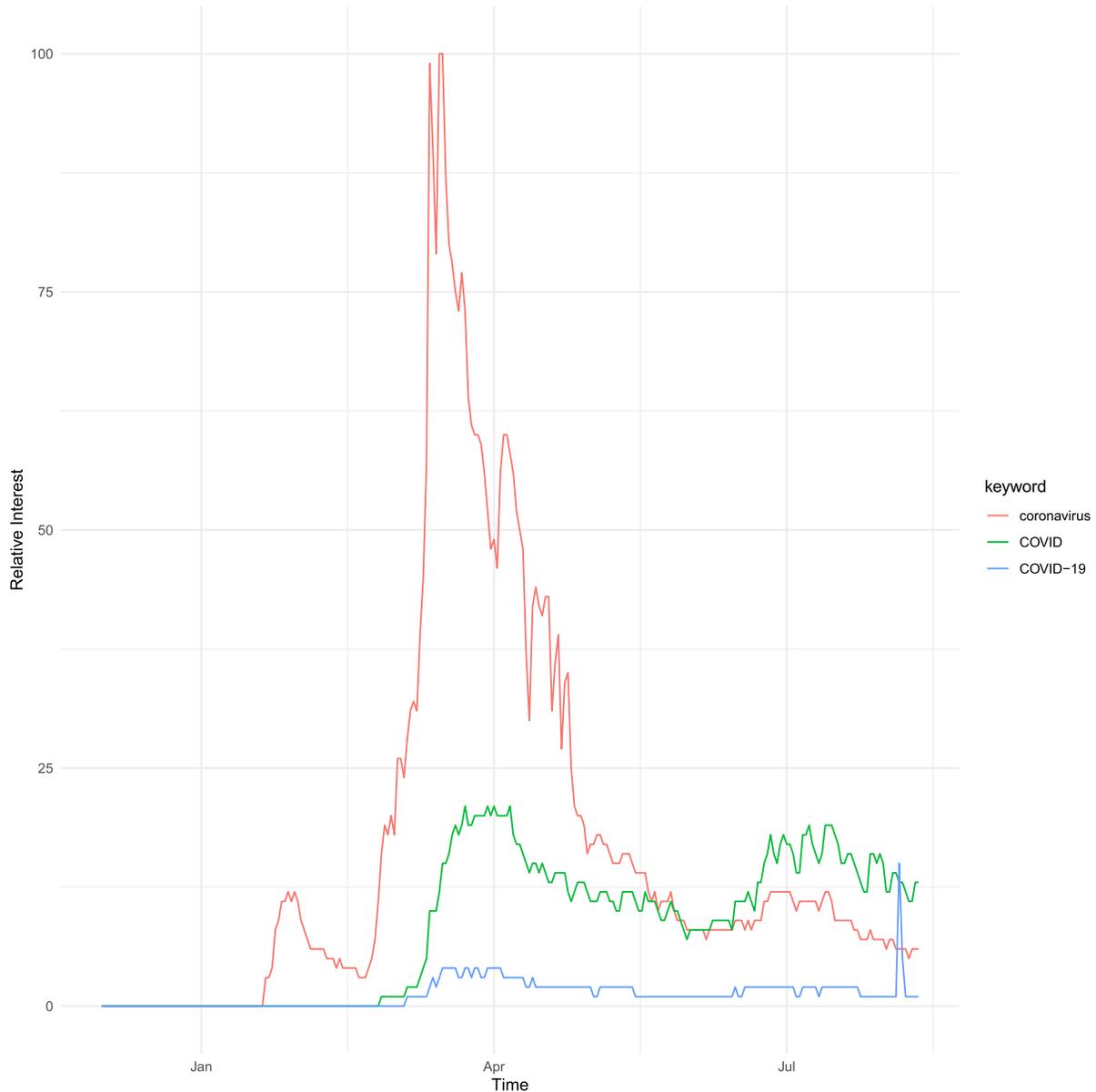


Figure 1: Visualization of Google trends web searches for “coronavirus,” “COVID,” and “COVID-19” from December 2019 to August 2020.

The corpus of Google searches is preprocessed by removing extraneous characters, stopwords (as well as the base search term “coronavirus” in order to focus on *related* terms), stripping whitespace, and making all terms lowercase.

Upon preprocessing, staging the text data from related Google searches includes three steps. First, we create a term-document matrix (TDM) which gives the frequencies of terms across Google searches. Then, we translate the TDM into a term-term matrix (TTM), which gives the frequencies of co-occurrence across all terms with each other. The TTM is a necessary step allowing for the final stage, which is to create the adjacency matrix,  $A_{ij}$ , which is a square matrix

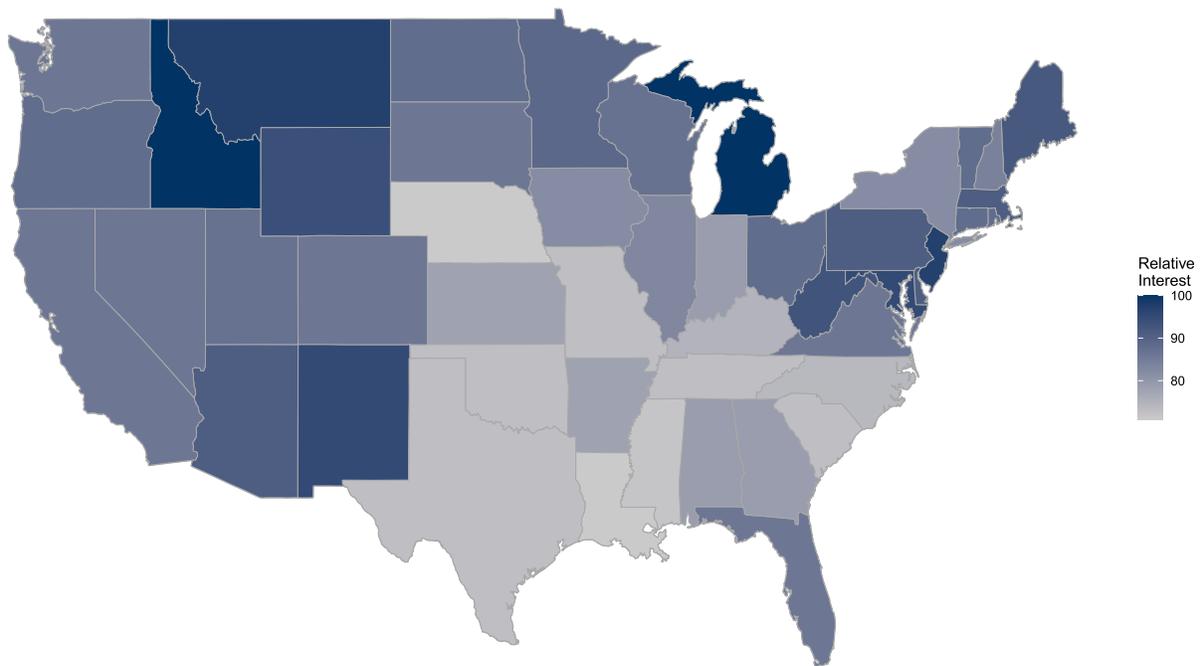


Figure 2: Visualization of relative interest in “Coronavirus” from Google trends searches by state across America from January 16, 2020 to August 11, 2020.

with elements as either 0 indicating no connection between vertices,  $v_i$  and  $v_j$ , or 1 indicating a connection between vertices (e.g., all elements on the main diagonal are of value 1). That is,

$$A_{ij} = \begin{cases} 1 & \text{if vertices } v_i \text{ and } v_j \text{ are connected,} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Based on the adjacency matrix, we constructed an undirected graph,  $G = (V, E)$ , on the full search space, where  $V$  contains the full set of vertices and  $E$  contains all the edges between vertex  $I$  and  $j$  such that  $A_I = 1$ . We also built individual networks for each month of Google searches to show the evolution of the structure of the Google search space over time. The month-based results are presented in the *Supplementary Material* in Figures 2–9. The full network shown in Figure 3, and is used in the final stage of analysis. And, the patterns of node degrees are displayed in Figure 1 in the *Supplementary Material*.

In the final stage, we fit and compare three widely used community detection algorithms to deepen the exploration of the Google search space. There are many ways to conceptualize community detection (Orman et al., 2012; Fortunato, 2010). For example, we could think of the formation of communities in a single space as a random walk, where modules of vertices in the network are recovered by taking brief, random walks. The algorithm for this approach is called “walktrap” (Pons and Latapy, 2005), and is premised on the assumption that the shorter the walks, the higher the likelihood of local structure, compared to longer walks meaning less local similarity across the network. Though there are many others, we proceed with three widely used approaches and compare results to strengthen reliability of the patterns. If patterns across all three are consistent, then this would be strong evidence that communities likely characterize

this search space. If results vary, then perhaps the search space is not well partitioned, meaning people self-educate on the Coronavirus in unique, non-constant ways. The three algorithms are: the Girvan-Newman algorithm (often called “edge betweenness”) (Girvan and Newman, 2002; Despalatović et al., 2014), propagating labels (Raghavan et al., 2007), and the Clauset-Newman-Moore algorithm (often called “greedy optimization of modularity”) (Clauset et al., 2004). Edge betweenness iteratively removes edges with the shortest number of paths through it, resulting in a rooted tree (dendrogram) structure with labels derived at different splits in the tree. The propagating labels algorithm is local, resembling a neighbor-based approach to clustering vertices, where labels are derived based on the majority of similar vertices in a small neighborhood. These labels, which are derived on the basis of local looks at the data, are iteratively updated and the algorithm stops when the partitioning of the space no longer changes. Greedy optimization of modularity operates by greedily searching the space to derive labels on the basis of the maximum modularity scores. Though a computationally efficient and widely used local method, some have discovered problems with module size and scaling (see, e.g., Fortunato and Barthelemy, 2007).

The value of community detection in this and similar applications is to offer more precise exploration as to whether local communities characterize some network space, which in our case is the Google search space related to “coronavirus.” If communities are uncovered, the network would be comprised of more densely connected modules or “clusters”, implying uniformity in self-education. This pattern is compared to the alternative of a sparse space, suggesting people are searching for and learning about coronavirus in very different ways, such that no structure emerges. The value of these three algorithms, and thus the justification for selecting them, is their widespread usage, understanding, and theoretical grounding. Understanding of the algorithms and the patterns that emerge, then, will be less burdensome, compared to other less-well known algorithms that may limit interpretability.

The first algorithm calculates edge betweenness as the shortest distance of any path traveling through the calculated edge. Girvan and Newman (2002) applied this definition to edges to locate the shortest paths between modules. By locating the shortest paths (suggesting similarities across vertices) compared to longer paths (indicating greater sparsity and thus less similarity), we are able to iteratively home in on a likely module/community. Through a process of progressively removing paths with the highest edge betweenness scores, the algorithm constructs a hierarchical tree (“dendrogram”), which represents the latent structure of the network.

The second “propagating labels” algorithm finds an optimal representation of community when the labels for all vertices,  $v \in V$ , have broken ties and thus look like a majority of the other labels surrounding the candidate vertex,  $v_i$ . Importantly, this algorithm, which is detailed in section three of Raghavan et al. (2007), shuffles the vertex labels to be in random order, and then attempts to recover a version of the network where smaller neighborhoods of vertices surrounded by a majority of like-labeled vertices. This strategy informs the labels assigned to vertices, which is the indication of the community structure.

Finally, for a different approach to community detection, the third algorithm is interested in finding structure of subgraphs within the network based on a maximal number of edges included in a community, compared to some expected value of a random version of the same network. This process of attempting to find structure from a randomized version of the data is indeed common and is similar to other techniques concerned with data representation, e.g., uniform manifold approximation and projection (McInnes et al., 2018). The formalization is not reproduced here

for threat of redundancy, but is clearly and simply laid out in “Equations (1)–(7)” in “Section 2” of Clauset et al. (2004).

In sum, all of these algorithms have a common goal: to search for and uncover latent, non-random structure in the network. Yet, though the goal is common, the process for uncovering this structure is quite different as previously discussed. The value of this empirical approach for present purposes is to offer three very differently constructed, but similarly motivated methods for a more holistic and thorough search of the Google search space.

The empirical strategy of this analysis is based on uncovering and understanding how people self-educate in the current global pandemic when they are more restricted in in-person interaction, and whether search patterns are more similar or different on average in these searching patterns.

### 3 Network of the Full Google Search Space

The network of the full Google search space related to “coronavirus” is shown in Figure 3. There are two notable trends that emerge. The first is that the full search space is partitioned into roughly two halves.

The first half includes a few dense regions around search terms like “cases,” “symptoms,” and “county.” These make sense at an intuitive level given people’s desires to know where new cases and outbreaks are occurring, and also in search of information on best to self-diagnose symptoms in line with work showing increased individual behavior online for similar goals (Collado-Borrell et al., 2020). These are indeed core aspects of self-education relating to a rapidly spreading virus like coronavirus in the context of a rapidly growing information environment (Gupta et al., 2020).

The other half of the search space though is much sparser, with very little connection across related searches. For example, there is a range of searches on distinct, nuanced topics like “Tom Hanks,” “student loan forgiveness,” “ibuprofen,” and “Ecuador.” This suggests that roughly half of the people Googling coronavirus are interested in more niche aspects often only tangentially related to the virus, but perhaps more directly related to personal contexts (“student loan forgiveness”) or interests (e.g., “Tom Hanks” in Figure 4 in the *Supplementary Material*). At this stage, it appears that the likely communities that characterize the former, denser part of the search space may be concentrated around more overtly related topics of broader interest.

### 4 Community Detection

Community detection is most useful to explore whether latent structure exists in a network. *Latent* and *structure* in this context translates to unobserved and uncoordinated (in that people are Googling in isolation), yet all the while self-educating in similar ways. Similarity in network structure manifests in the form of modules or “clusters.” That which defines a module is twofold: the density of connections within a module, and also the sparsity of connections between other modules. If the former and the latter are high, then the space is considered to have high modularity, substantively suggesting a clear partitioning of the Google search space. Our goal in this final section, then, is to explore whether latent structure exists in this space or not, which will push us closer to understanding how people self-educate in a global pandemic of this sort.



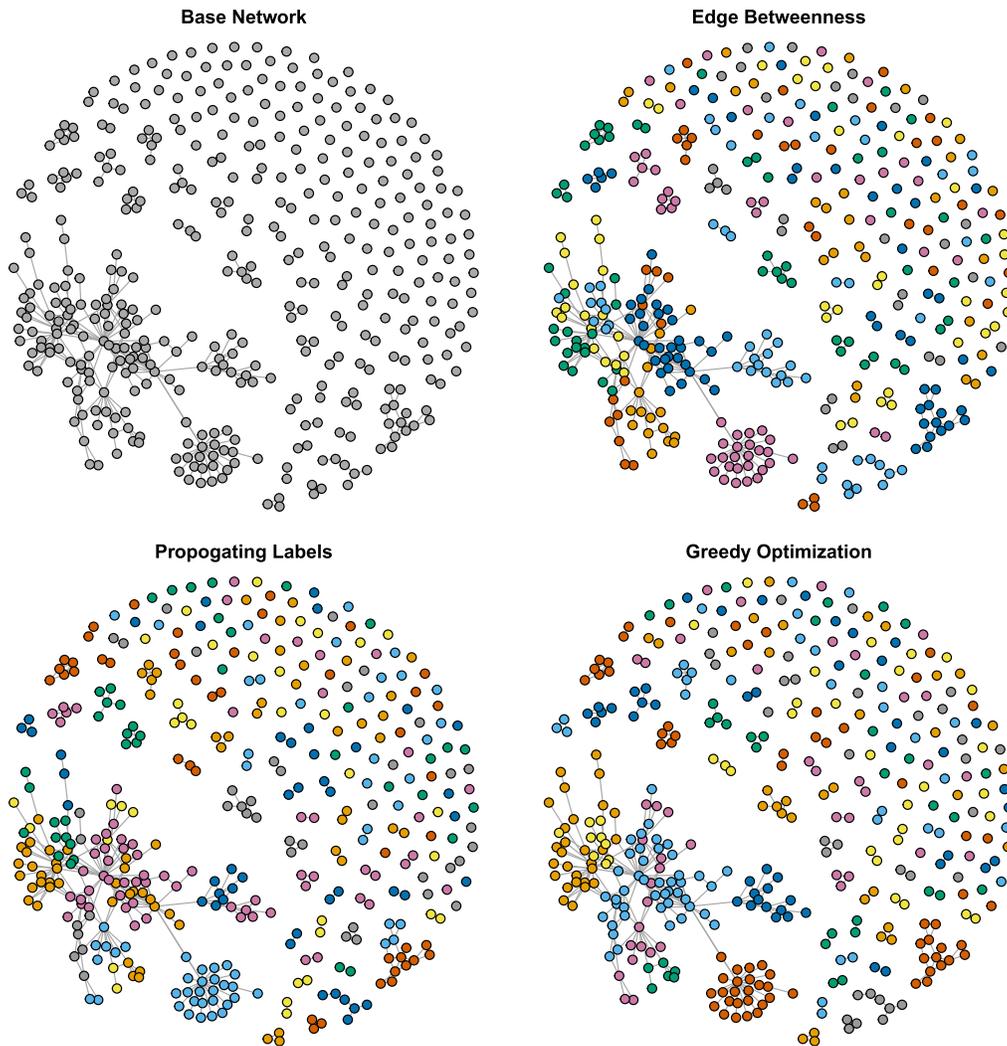


Figure 4: Results comparing three community detection algorithms used to explore the Google search space.

analysis is an exploratory first look at the Google search space, no such constraint was employed to allow for natural structure to emerge.

In addition to the sparsity of part of the network, all three algorithms detected some consistent and dense trends, including a community surrounding the search term “county.” Further, there is similarity in other densely connected regions surrounding “cases,” “symptoms,” and “stimulus check.” The similarities across the algorithms relating to these key terms are important, because regardless of the specific algorithm, we would expect clear, non-random communities to emerge if such regions are truly present in the data. As these regions are indeed clear and consistent across the algorithms, the structure of this Google search space is at least in part indeed non-random. This is the benefit of selecting three different, but widely understood algorithms to detect community pattern.

Yet, despite the similarities, there is also variance in the labels found by the algorithms across these communities. For example, the propagating labels algorithm treated the “county”

community similarly to the “update” and “symptoms” communities, though these few communities remain largely stable across each algorithm. Interestingly, the dense communities partitioning varied across all algorithms, with surprisingly few communities found in the greedy optimization of modularity algorithm in the lower right plot. This is surprising because of the algorithm’s focus on locally and greedily searching for communities, rather than a broader and more global focus. A naive expectation, thus, might be more fragmentation in communities from greedy optimization, but this is not what is seen in Figure 4.

In sum, these algorithms are all picking up on Googling interest in geolocations of COVID-19 outbreaks, in line with work showing such variation is fast moving and widespread at both the U.S. county level (Javan et al., 2020), and abroad (Pobiruchin et al., 2020). And from a technical perspective, the value of fitting multiple algorithms to a common data space such as in our case, is that different versions of community construction are able to emerge, which ultimately offer the researcher greater flexibility in interpretation. Such a step is valuable in an exploratory study of this sort to pave the way for future, more targeted and causal studies.

## 5 Concluding Remarks

There are a few broad conclusions to draw based on these results. First, people tend to search in an uncoordinated, though still common fashion relating to more obvious terms associated with coronavirus (e.g., “location,” “cases,” “symptoms”). Yet, for much of the space (and over time), there remain shifting and tangential-type queries that vary widely, suggesting the space is not clearly partitioned.

These two themes make intuitive sense as a start. Where the information environment surrounding the Coronavirus is both in flux with new details emerging frequently but also stable at a higher level regarding common themes like outbreak locations and symptoms, the individual search space takes on a unique character. About half of the search space includes a few prominent communities related to more common and understandable topics previously described, whereas the other half of the search space is less stable, and more rapidly evolving.

There are a number of future studies that could build on these exploratory findings. First, time is a likely driver of structure of this space given the rapidly shifting information environment referenced throughout. For example, do key exogenous events such as the U.S. electoral cycle prompt structural shifts in search patterns or not? Further, future work could pair these search data with social media data to explore whether parallel trends exist. Namely, do we see a similar type of split between a more stable, but narrower battery of search terms/topics paired with a wide array of tangential, shifting search terms/topics? Finally, anomaly detection would be useful to ask similar questions, but with more data as we live with the Coronavirus longer. For example, might these patterns be “seasonal,” or could they be anomalous, where people’s attention moves away from these types of issues over time?

Ultimately, beyond providing a deeper understanding of the contours of self-education in this unprecedented season of the Coronavirus and COVID-19, this research corroborates the base understanding of the extreme value of Google in modern society as a tool for self-education, which is a critical skill in a time of isolation and fear.

## Supplementary Material

The following are included in online Supplementary Material:

- Network Node Degrees (Full Network).
- Network of Google Search Trends: January.
- Network of Google Search Trends: February.
- Network of Google Search Trends: March.
- Network of Google Search Trends: April.
- Network of Google Search Trends: May.
- Network of Google Search Trends: June.
- Network of Google Search Trends: July.
- Network of Google Search Trends: August.
- Replication R code.

## References

- Aljilani Y, Kadobayashi R (2015). Games and their overlooked potential: Facilitating learning using entertainment games. In: *2015 IIAI 4th International Congress on Advanced Applied Informatics*, 291–296. IEEE.
- Anastassopoulou C, Russo L, Tsakris A, Siettos C (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS ONE*, 15(3): e0230405.
- Apuke OD, Omar B (2020). Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, 56: 101475.
- Atlas S, Porto N, Jian Xiao J (2018). Financial education and confidence in financial knowledge. In: *NA – Advances in Consumer Research* (A Gershoff, R Kozinets, T White, eds.), volume 46, 23–27. Association for Consumer Research, Duluth, MN.
- Bastani P, Bahrami MA (2020). COVID-19 related misinformation on social media: A qualitative study from Iran. <https://doi.org/10.2196/18932>.
- Baud D, Qi X, Nielsen-Saines K, Musso D, Pomar L, Favre G (2020). Real estimates of mortality following COVID-19 infection. *Lancet. Infectious Diseases*. 20(7): 773.
- Cairney P, Wellstead A (2020). COVID-19: Effective policymaking depends on trust in experts, politicians, and the public. *Policy Design and Practice*. Forthcoming, <https://doi.org/10.1080/25741292.2020.1837466>
- Chan AK, Nickson C, Rudolph J, Lee A, Joynt G (2020). Social media for rapid knowledge dissemination: early experience from the COVID-19 pandemic. *Anaesthesia*, 75(12): 1579–1582.
- Clauset A, Newman ME, Moore C (2004). Finding community structure in very large networks. *Physical Review E*, 70(6): 066111.
- Collado-Borrell R, Escudero-Vilaplana V, Villanueva-Bueno C, Herranz-Alonso A, Sanjurjo-Saez M (2020). Features and functionalities of smartphone apps related to COVID-19: Systematic search in app stores and content analysis. *Journal of Medical Internet Research*, 22(8): e20334.
- Croce MMM, Farroni P, Wolfskeil I (2020). When the markets get COVID: Contagion, viruses, and information diffusion. [https://cepr.org/active/publications/discussion\\_papers/dp.php?dpno=14674](https://cepr.org/active/publications/discussion_papers/dp.php?dpno=14674).
- Despalatović L, Vojković T, Vukicevic D (2014). Community structure in networks: Girvan-Newman algorithm improvement. In: *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 997–1002. IEEE.

- Dinh L, Parulian N (2020). COVID-19 pandemic and information diffusion analysis on Twitter. *Proceedings of the Association for Information Science and Technology*, 57(1): e252.
- Dixit A, Marthoenis M, Arafat SY, Sharma P, Kar SK (2020). Binge watching behavior during COVID 19 pandemic: A cross-sectional, cross-national online survey. *Psychiatry Research*, 289: 113089.
- Effenberger M, Kronbichler A, Shin JI, Mayer G, Tilg H, Perco P (2020). Association of the COVID-19 pandemic with internet search volumes: A Google Trends™ analysis. *International Journal of Infectious Diseases*, 95: 192–197.
- Elmer T, Mepham K, Stadtfeld C (2020). Students under lockdown: Comparisons of students’ social networks and mental health before and during the COVID-19 crisis in Switzerland. *PLoS ONE*, 15(7): e0236337. <https://doi.org/10.1371/journal.pone.0236337>.
- Fortunato S (2010). Community detection in graphs. *Physics Reports*, 486(3–5): 75–174.
- Fortunato S, Barthelemy M (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1): 36–41.
- Fortunato S, Hric D (2016). Community detection in networks: A user guide. *Physics Reports*, 659: 1–44.
- Fruchterman TM, Reingold EM (1991). Graph drawing by force-directed placement. *Software, Practice & Experience*, 21(11): 1129–1164.
- Girvan M, Newman ME (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12): 7821–7826.
- Grasselli G, Pesenti A, Cecconi M (2020). Critical care utilization for the covid-19 outbreak in Lombardy, Italy: Early experience and forecast during an emergency response. *JAMA*, 323(16): 1545–1546.
- Grech V (2020). Unknown unknowns – COVID-19 and potential global mortality. *Early Human Development*, 144: 105026.
- Gunnell D, Appleby L, Arensman E, Hawton K, John A, Kapur N, et al. (2020). Suicide risk and prevention during the COVID-19 pandemic. *The Lancet Psychiatry*, 7(6): 468–471.
- Gupta L, Gasparyan AY, Misra DP, Agarwal V, Zimba O, Yessirkepov M (2020). Information and misinformation on COVID-19: a cross-sectional survey study. *Journal of Korean Medical Science*, 35(27): e256.
- Hartley K, Jarvis DS (2020). Policymaking in a low-trust state: Legitimacy, state capacity, and responses to COVID-19 in Hong Kong. *Policy and Society*, 39(3): 403–423.
- Javan E, Fox SJ, Meyers LA (2020). Probability of current COVID-19 outbreaks in all us counties. *The University of Texas at Austin Technical Report*. [https://cid.utexas.edu/sites/default/files/cid/files/covid-risk-maps\\_counties\\_4](https://cid.utexas.edu/sites/default/files/cid/files/covid-risk-maps_counties_4), 3.
- Kim SJ, Bostwick W (2020). Social vulnerability and racial inequality in COVID-19 deaths in Chicago. *Health Education & Behavior*, 47(4): 509–513.
- Laato S, Islam AN, Farooq A, Dhir A (2020). Unusual purchasing behavior during the early stages of the COVID-19 pandemic: The stimulus-organism-response approach. *Journal of Retailing and Consumer Services*, 57: 102224.
- Lehmiller JJ, Garcia JR, Gesselman AN, Mark KP (2020). Less sex, but more sexual diversity: Changes in sexual behavior during the COVID-19 coronavirus pandemic. *Leisure Sciences*. Forthcoming, <https://doi.org/10.1080/01490400.2020.1774016>
- Liu Q, Zheng Z, Zheng J, Chen Q, Liu G, Chen S, et al. (2020). Health communication through news media during the early stage of the COVID-19 outbreak in China: Digital topic modeling approach. *Journal of Medical Internet Research*, 22(4): e19118.

- Ma R, Deng Z, Wu M (2020). Effects of health information dissemination on user follows and likes during COVID-19 outbreak in China: Data and content analysis. *International Journal of Environmental Research and Public Health*, 17(14): 5081.
- McInnes L, Healy J, Melville J (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint: <https://arxiv.org/abs/1802.03426>.
- Meier K, Glatz T, Guijt MC, Piccininni M, Van Der Meulen M, Atmar K, et al. (2020). Public perspectives on protective measures during the COVID-19 pandemic in the Netherlands, Germany and Italy: A survey study. *PLoS ONE*, 15(8): e0236917.
- Orman GK, Labatut V, Cherifi H (2012). Comparative evaluation of community detection algorithms: A topological approach. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(08): P08001.
- Pennycook G, McPhetres J, Zhang Y, Lu JG, Rand DG (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7): 770–780.
- Perc M, Gorišek Miksić N, Slavinec M, Stožer A (2020). Forecasting COVID-19. *Frontiers in Physics*, 8: 127.
- Pobiruchin M, Zowalla R, Wiesner M (2020). Temporal and location variations, and link categories for the dissemination of COVID-19-related information on Twitter during the SARS-COV-2 outbreak in Europe: Infoveillance study. *Journal of Medical Internet Research*, 22(8): e19629.
- Pons P, Latapy M (2005). Computing communities in large networks using random walks. In: *International Symposium on Computer and Information Sciences* (P Yolum, T Güngör, F Gürgeç, C Özturan, eds.), 284–293. Springer.
- Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, et al. (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *The Lancet Public Health*, 5(5): e261–e270.
- Raghavan UN, Albert R, Kumara S (2007). Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 76(3): 036106.
- Rovetta A, Bhagavathula AS (2020). COVID-19-related web search behaviors and infodemic attitudes in Italy: Infodemiological study. *JMIR Public Health and Surveillance*, 6(2): e19374.
- Saltzman LY, Hansel TC, Bordnick PS (2020). Loneliness, isolation, and social support factors in post-covid-19 mental health. *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(S1): S55–S57.
- Sher L (2020). The impact of the COVID-19 pandemic on suicide rates. *QJM: An International Journal of Medicine*, 113(10): 707–712.
- Song P, Karako T (2020). COVID-19: Real-time dissemination of scientific information to fight a public health emergency of international concern. *Bioscience Trends*, 14(1): 1–2.
- Tuite Bogoch AR II, Sherbo R, Watts A, Fisman D, Khan K (2020). Estimation of coronavirus disease 2019 (COVID-19) burden and potential for international dissemination of infection from Iran. *Annals of Internal Medicine*, 172(10): 699–701.
- van der Linden S, Roozenbeek J, Compton J (2020). Inoculating against fake news about COVID-19. *Frontiers in Psychology*, 11: 2928.
- Venkatesh A, Edirappuli S (2020). Social distancing in COVID-19: What are the mental health implications? *BMJ*, 368: m1089.
- Waggoner PD (2020). Pandemic policymaking: Learning the lower dimensional manifold of congressional responsiveness. arXiv preprint: <https://arxiv.org/abs/2011.04763>.

Ward A, Grillo TLH, Fernbach PM (2018). From novice to know-it-all: How Google-based financial learning affects financial confidence and decisions. In: *NA – Advances in Consumer Research* (A Gershoff, R Kozinets, T White, eds.), volume 46, 23–27. ACR North American Advances, Duluth.