Does the Bubble Go Beyond?

An Exploration of the Urban Filter Bubble

Annelien Smets
imec-SMIT, Vrije Universiteit Brussel
Belgium

Eladio Montero
imec-SMIT, Vrije Universiteit Brussel
Belgium

Pieter Ballon
imec-SMIT, Vrije Universiteit Brussel
Belgium

ABSTRACT

The increasing prevalence of algorithms in our everyday lives has raised concerns about their societal effects. Algorithmic personalization is said to create filter bubbles that threaten democracy by curating the content users are exposed to. However, with our urban environments becoming increasingly digitally layered, they become scope of algorithmic curation as well. We therefore argue that the urban context should also be scope of algorithmic impact assessments to avoid the creation of urban filter bubbles; people only being exposed to a specific part of the city, which differs from what others see because of algorithmic personalization. In this paper, we present a methodology to assess the urban filter bubble hypothesis and perform a preliminary study to verify our approach.

CCS CONCEPTS

• **Information systems** → *Personalization.*

KEYWORDS

algorithmic curation, diversity, location, urban filter bubble

1 INTRODUCTION

There has been a fundamental shift in the way we consume information. Not only is there a change in the kind of media through which we acquire information, these new intermediaries are also characterized by evoking information overload. This, in turn, triggers innovations that automatically filter selections of information resulting in "more and more of the information we receive in the world [being] curated by algorithms" [17]. This algorithmic curation of information is said to be particularly present in online search engines, personalizing their search results based on an individual's characteristics [9]. However, this kind of personalization also characterizes recommender systems for music, movies, products to buy, job offers, and even our potential partners. Yet, despite its omnipresence, it has raised concerns about *filter bubbles*; the idea that information diversity is diminishing and users are only exposed to information they agree with [19]. Despite the fact that empirical evidence for the filter bubble is sparse [9, 16, 17], we acknowledge the relevance of the debate. Algorithms are increasingly becoming part of our everyday lives and cause scholarship to ask critical questions about their societal impact [7, 25].

With our urban environments becoming increasingly digitally layered, the urban context is no longer excluded from algorithms and should be scope of critical assessments as well. In this work, we therefore argue that the filter bubble debate should go beyond, and also take into account the potential consequences in a physical context. To the best of our knowledge, this study is the first attempt to discuss the concept of an urban filter bubble. The aim of this paper is therefore two-fold: (1) demonstrate the need to consider the consequences of algorithmic curation in the urban context, and, more specifically, (2) discuss a methodological framework to assess the potential reach of the urban filter bubble.

It is important to note that this study does not aim to assess the existence of an urban filter bubble, rather the objective is to get insights on our proposed methodology and defined measures. This paper will therefore mainly discuss results that are contributing to this objective. In the next section, we elaborate on the concept of an urban filter bubble and its relation to existing work. Subsequently, we propose our methodology to study the urban filter bubble, and discuss our pilot experiment. Based on the latter, we conclude this paper by presenting our lessons learned and avenues for future work.

2 THE URBAN FILTER BUBBLE

In this section we elaborate the filter bubble concept by extending it to the urban sphere and discuss how it relates to previous work.

2.1 Conceptualization

One of the main arguments in the filter bubble debate is algorithmic curation, indicating that content exposure is no longer the result of a human selection process [9]. With the urban environment being pervaded by IoT technologies, it becomes susceptible to algorithmic curation as well. However, the nature of these technologies and their inherent connection with the physical environment suggest that we are not talking about content exposure but rather context exposure.

The way we are exposed to the urban space is no longer a mere result of decisions by urban planners and architects. The fact that our cities are becoming digitally layered urban environments [22] sets the scene for applications such as Waze or Google Maps to be an additional context curator by guiding us through the urban space. These applications have not only become our primary guide to determine *how* we navigate through the urban environment, they are becoming a dominant source to decide *where* we are going as well. Research on travel information search shows that we are increasingly turning to online information for destination decision-making [11] and the rise of social media has even amplified this behavior [26]. The objective of online applications to become context curators is also exemplified by Google Maps' recent feature which recommends places to go to [3]. Obviously, the mere availability and use of information about the urban environment does not necessarily imply that our choices are being algorithmically steered by these applications. Except that there is more: Pan et al. [18] showed that users are biased towards higher ranked results

ImpactRS '19, September, 2019, Copenhagen, Denmark

even when there are doubts about the relevance of the search result. This implies that when someone uses a search engine to look for a particular thing to do in the city, it is likely that they will only consider the first results. Although this argument can be contested by appealing to the autonomy of the human decision-maker, it becomes essential when thinking of algorithmic decision-makers. It is no longer science-fiction to think of a scenario where we ask our digital personal assistant to make a dinner reservation [14] and if we do not specify the venue ourselves, it will be algorithmically picked. What would this algorithmic decision-making mean for our explorative right to the city?

This example demonstrates the need to consider the consequences of algorithmic curation in the physical sphere and eventually the curation of our experiences as well. Especially in urban environments, cities, where diversity and the right to the city are fundamental aspects we should take into account the contextual consequences of algorithmic curation. After all, the consequences go beyond the mere experience of citizens, and there are also political and economic interests at stake. If an algorithm is deciding how people are navigating through a city, which parts of the city they are exposed to; it is deciding where they spend their time, where they spend their money.

In this study, we therefore explore the filter bubble hypothesis in an urban context. We start from the assumption that people are increasingly using online search engines to get recommendations on what to do in the city or where to go. The urban filter bubble then articulates the idea that by using the results of online search engines people are likely to only be exposed to a specific part of the city, which differs from what others see because of algorithmic personalization.

2.2 Related work

The filter bubble hypothesis has been mainly studied within the domain of public policy discourse, where it is linked to concepts such as echo chambers [6, 23] and viewpoint diversity [4]. While those studies focus on the impact of personalization, there is a vast amount of research discussing the actual techniques for algorithmic search personalization [5] and some studies specifically focus on geolocation [1, 27]. Most of the work described in this paper builds upon the research of Kliman-Silver et al. [13], showing that differences in search results grow as physical distance increases. However, in this work, we take the online search results back to their physical location in the urban space and question what personalization might imply for the places we actually visit in the city.

3 METHODS

In this section, we describe our search methodology, followed by outlining the different diversity measures that are used to analyze the results. Finally, we discuss the research design of the pilot study to validate our approach.

3.1 Search methodology

Considering both Google's and Google Maps' high market share [21] this study focuses on search results from Google Maps. We specifically select Google Maps' search results as they are guaranteed to have a corresponding geolocation.

The queries in Google Maps are performed manually from a blank state computer with no cookies stored, using Google Chrome. To vary the search location (see 3.3) the GPS location sensors on Google Chrome are overridden. This option was chosen instead of changing the IP address location with a VPN, because "Google Search personalizes search results largely based on the provided GPS coordinates rather than the IP address" [13].

The construction of the search terms is based on a set of travel-related terms that are found to be frequently used when exploring a new destination [26]. To account for differences that might be related to proximity [13], we combine these search terms with 3 cities (Brussels, Ghent and Berlin) that have a different proximity to our search location. This results in the following 9 queries:

- Restaurant in <Brussels/Ghent/Berlin>
- Hotel in <Brussels/Ghent/Berlin>
- What to do in <Brussels/Ghent/Berlin>

After performing the queries, we scrape the results using the lmxl package in Python [15]. This way, each element can be mapped using an XPath (XML Path Language) route [24].

3.2 Diversity measures

Spending time in a physical environment implies that we are doing *something somewhere*. To account for these two concepts, we expand the diversity measurement from similar studies [10] to a two-dimensional concept. We use two measures of diversity based on respectively **content** and **context**. The former is similar to other diversity measures in related studies [10], while the latter is based on the actual location coordinates of the search results and expresses how they are distributed among the urban space. Moreover, these diversity measures are calculated for two perspectives: the **intra-diversity** perspective accounts for the diversity among one's individual search results (how diverse are my search results?), while the **inter-diversity** perspective considers the diversity among the search results of multiple users (how much do my search results differ from yours?). This 2x2 framework results in a set of diversity measures and their corresponding calculations as shown in Table 1.

The **purity** accounts for the diversity among one's individual search results by calculating the extent to which the result set (i.e. cluster) contains elements of a single class.

The **Jaccard-index** defines the overlap in search results between two sets: 0 when there is no overlap, and 1 when they have exactly the same results. This index does not include the order of the results and therefore the **edit distance** is taken into account as well: it shows how many transformations (i.e. insert, delete, substitute) are needed to make the lists identical.

The **total sum of squares** represents the sum of all pairwise square distances between the locations of the search results. This is an indication of how disperse the results are.

The **cluster centroid distance** points at the distance between the centroids of two result sets and can be used to assess if different users are indeed seeing different parts of the city. The cluster centroid is calculated as the average of the location coordinates of the elements in the cluster (i.e. the set of search results).

Table 1: Diversity measures

	Intra-diversity	Inter-diversity
Content	Purity	Jaccard-index, edit distance
Context	Cluster total sum of squares	Cluster centroid distance

3.3 Study design

We set up a small pilot study to validate our methodology and assess the value of our defined metrics. To this end, we use an agent-based testing approach [9] and define 3 features that will be used for personalization: language, search location and user profile.

The **language** of the queries varies between Dutch and French, two of the official languages in Belgium. We chose to vary the language of the search queries because it is a realistic reflection of the Belgian population.

The next variable is the **search location**, i.e. the location from which the searches are being performed. We query from two distinct locations in the Brussels area, based on the idea that "users' geolocation can be used as a proxy for other demographic traits" [13]. The search locations are Sint-Jans-Molenbeek (SJM) and Sint-Pieters-Woluwe (SPW), two municipalities in Brussels that show significant differences in terms of unemployment rate and education levels, which are socio-demographic traits that are found to have an impact on the kind of activities people do in a city [2].

Finally, we construct two different **user profiles** that are used to perform the queries. These profiles vary mainly in terms of educational level, employment status and family situation as these have been indicated as influencing the kind of cultural activities one undertakes. Similar to previous studies [8–10], we build the user profiles by means of two phases. In the training phase, we feed the profile by (1) searching Google for five agent-specific terms, (2) adding a related product in an online shopping cart and (3) browsing through 5 articles in a specified media outlet. We execute this training phase for five consecutive days and alternate the training phase with the testing phase during which the search queries are performed. This is in contrast to the above-mentioned studies, who did not alternate these phases. Our decision to alternate the phases is motivated by the goal to observe the evolution of the personalization over time. Using this approach, we build two user profiles which we refer to as Agent A and Agent B.

4 RESULTS

Based on the aforementioned methodology we collected 4,988 search results that consist of 422 unique locations.

4.1 Content diversity

The Jaccard-index and edit distance have been used in similar research to assess content diversity [10, 13], and prove to be valuable in our pilot study as well. For example, Figure 1 shows the average Jaccard-indexes for the Dutch and French search results (e.g. *Hotel in Brussel* and *Hôtel à Bruxelles*). Over the 9 queries, the average Jaccard-index is 0.31, indicating that on average 31% of results appear in both French and Dutch search results. In contrast to what we would expect based on search proximity, Figure 1 indicates that the overlap does not increase when the physical distance increases. The average edit distance is 23 and taking into account that there are approximately 20 results in each list and there is only 30% similarity, this edit distance indicates that the order of the shared results is quite similar.

The Jaccard-indexes for the categories of the *what to do* queries are also represented in Figure 1. The similarity of the categories is on average 23.7%, which means that the different search results not only represent different locations but also different kinds of activities. Nevertheless, the average purity[1] of the Dutch and French *what to do* results is respectively 0.64 and 0.66. Hence, within one language, on average 65% of the results are of the same category and neither language shows more diversity in these categories.

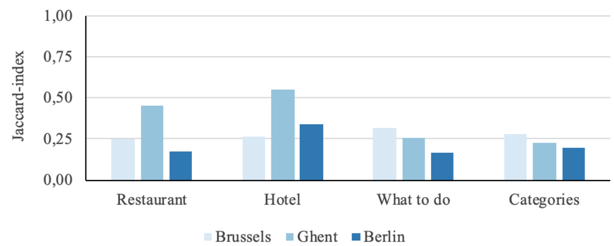


Figure 1: Average Jaccard-index: Dutch vs French.

On the other hand, the diversity among the results of queries from two search locations (SJM and SPW) is significantly smaller. The average overlap between those search results is 77.2%, more than twice the overlap of the French and Dutch search results. Hence, in this case, the personalization due to language is much stronger compared to the one based on search location. In line with previous research [13], the Jaccard-indexes in Figure 2 illustrate that queries related to Brussels are more diverse than queries related to Ghent and Berlin. In this case, the search location (SJM or SPW) thus stronger influences the personalization when the proximity increases.

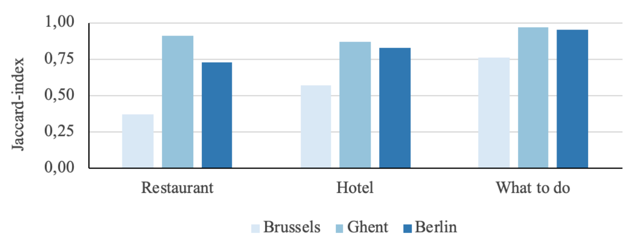


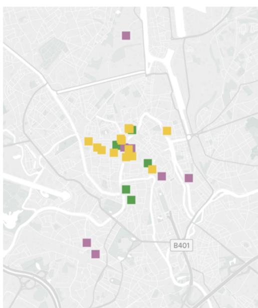
Figure 2: Average Jaccard-index: search locations.

[1] Purity was only calculated for the *what to do* results, since we only considered the most general categories. The categories of the *restaurant* and *hotel* results were consequently the same due to the nature of the query.

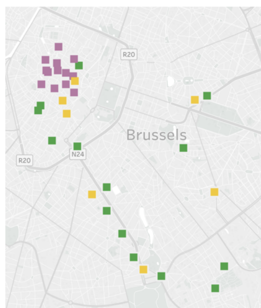
4.2 Context diversity

While the content diversity measures provide valuable and straightforward insights, the context measurements need some additional attention during interpretation.

For example, a visual inspection of the data in Figure 3(a) shows that generally most locations are clustered in the city center, with a few outliers for the Dutch results. However, our metric to indicate the dispersity of results (total sum of squares) does not account for outliers and consequently indicates that French results are 87% denser than Dutch results (see Figure 4(a) restaurants in Ghent). The mere interpretation of the measurement without the visual inspection could lead to a conclusion that is inconsistent with the actual data.



(a) Ghent



(b) Brussels

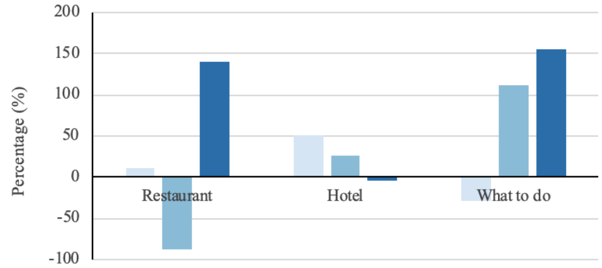
Figure 3: Restaurant results in Ghent and Brussels. Colors represent the query language in which the result appears: purple (only Dutch), green (only French) and yellow (both).

Although the same consideration in terms of outliers applies to the cluster centroid distance, this metric holds some potential as well. For example, as illustrated in Figure 4(b) the cluster centroid distance for restaurants in Ghent is 257, and 450 for Brussels. Indeed, Figure 3 shows that the French and Dutch search results are located closer to each other in Ghent (a) compared to Brussels (b).

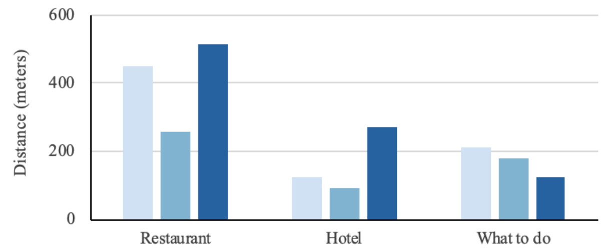
Finally, our preliminary results also demonstrate the need to account for both content and context diversity since one does not necessarily imply the other. For example, Figure 5 shows that French and Dutch results are practically covering the same parts of the Brussels' area (low context diversity), while the Jaccard-index of this result set is only 32% (high content diversity).

4.3 A note on user profiles

The reader might have noticed that we did not discuss any of the results related to the personalization based on user profiles. As explained before, we alternated the training phase with the testing phase to study the influence of the profile on the search results. Our assumption is that the more Google learns about the agent, the more personalization of search results would occur. Our main interest then lies in the geographic location of these search results. However, as Figure 6 shows, the overlap in the search results is not significantly decreasing, and, in some cases, even increases over time. The figures for other queries and agents are similar and we omit them for brevity. Despite the fact that Figure 6 shows an



(a) Relative density



(b) Centroid distance

Figure 4: (a) Relative density between Dutch and French search results. (b) Centroid distance between Dutch and French search results.

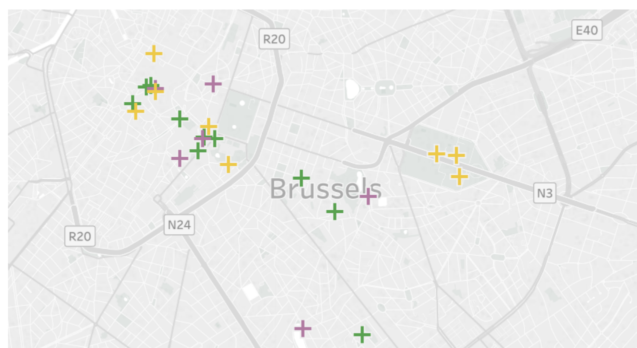


Figure 5: What to do results in Brussels. Colors represent the query language in which the result appears (see Figure 3).

interesting pattern for the Jaccard-index of Berlin, a closer look at the data did not provide any additional path to continue on. Additionally, the visual inspection of the data did not show significant differences in terms of locations of the search results, even not for those related to Berlin. Therefore, we will not continue the discussion of the diversity of these results, because it would lead beyond the scope of this paper.

5 CONCLUSION

This paper aims to contribute to the debate on algorithmic curation by addressing this topic in the urban context. To this end, we

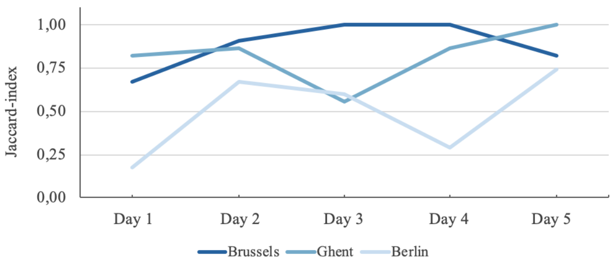


Figure 6: Jaccard-index: Agent A vs. Agent B (restaurants).

proposed a methodology that conceptualized diversity in a two dimensional way: content-context, and individual-group. Making these distinctions allows to (1) not only discuss the content diversity but also consider the physical diversity and (2) account for both individual diversity as well as diversity among multiple users. We argue that this multidimensional concept is required to account for the complexity of establishing diversity in an urban context and allows to foresee different configurations depending on the situation.

The results of our pilot study demonstrate the need to account for both content and context diversity, however, there is still room for improvement to our approach. Although content diversity can be assessed in a meaningful and straightforward way, the context measurements appear to require some adjustments to deal with the skewing impact of outliers in the physical dimension. Moreover, due to our manual approach, our study remains small scale while for an in-depth analysis of the urban filter bubble hypothesis it would be beneficial to collect data on a large scale. This would allow applying statistical methods to verify the hypothesis and implementing more variation, for example in terms of search locations or languages. Another limitation of our current study is that we have mainly focused on the context: where are the suggestions physically located. This focus arose due to our plea for the inclusion of context diversity. We did look at the content to a certain extent by taking into account the categories of the activities, however, future work could focus on studying this in more depth (e.g. price range). Finally, there is still the question of validity. One may indeed question if an agent-based testing approach could be said to be representative for an actual user. Nevertheless, taking into account the limitations of this pilot study and our recommendations, future studies could apply our approach to large scale data collected from real-world users.

The question of how to research the (societal) impact of algorithms is still part of an ongoing debate. However, we agree with authors like Kitchin [12] and Seaver [20] arguing that algorithms should be considered as algorithmic systems. Consequently, researching their impact is not limited to one methodology or assessment, but should encompass multiple research acts of which the work in this paper could be one. In line with this reasoning, we conclude this paper by asking questions that go beyond the mere issue of *measuring* the impact of algorithmic curation in the urban sphere, as they provide directions for future work. After all, we acknowledge that avoiding urban filter bubbles is not a straightforward exercise since the urban environment itself is already to a great extent characterized by districts or clusters. For example, many cities are known for having a fashion district, or areas that are populated by specific subgroups. Hence, there is a thin line between these inherent clusters and urban filter bubbles induced by algorithmic curation. These considerations stress the importance to critically assess the algorithmic curation of our contexts. In our understanding, this also comes with a thorough reflection on its societal impact including questions such as what is the best way to get people engaged with the city? Why is it beneficial to have an unfiltered discovery of the city? If there is an urban filter bubble, who is responsible to burst it?

In our future work, we will address some of these questions by exploring the value and meaning of serendipity in urban recommender systems. We hope that this work contributes to the debate of algorithmic curation and inspires others to continue on this topic.

ACKNOWLEDGMENTS

We gratefully acknowledge the help of Ms. Azadasl in collecting the data, and thank the anonymous reviewers for their constructive comments on an earlier version of this paper.

REFERENCES

[1] Leonardo Andrade and Mário J Silva. 2006. Relevance Ranking for Geographic IR.. In *GIR*.

[2] Julie Badisco, Ignace Glorieux, Lindsay Jacobs, and Katrien Lauwerysen. 2008. De weg naar cultuur. Ruimtelijke aspecten van culturele interesse en deelname aan het cultuuraanbod. (2008).

[3] Muhammad Ahmad Bashir, Sajjad Arshad, and Christo Wilson. 2016. "Recommended For You": A First Look at Content Recommendation Networks. In *Proceedings of the 2016 ACM on Internet Measurement Conference - IMC '16*. ACM Press, Santa Monica, California, USA, 17–24. https://doi.org/10.1145/2987443.2987469

[4] Engin Bozdag. 2015. *Bursting the filter bubble: Democracy, design, and ethics.* Ph.D. Dissertation.

[5] Zhicheng Dou, Ruihua Song, and Ji-Rong Wen. 2007. A large-scale evaluation and analysis of personalized search strategies. In *Proceedings of the 16th international conference on World Wide Web*. ACM, 581–590.

[6] Elizabeth Dubois and Grant Blank. 2018. The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society* 21, 5 (May 2018), 729–745. https://doi.org/10.1080/1369118X.2018.1428656

[7] Tarleton Gillespie. 2017. Algorithmically recognizable: Santorum's Google problem, and Google's Santorum problem. *Information, Communication & Society* 20, 1 (Jan. 2017), 63–80. https://doi.org/10.1080/1369118X.2016.1199721

[8] Mario Haim, Florian Arendt, and Sebastian Scherr. 2017. Abyss or Shelter? On the Relevance of Web Search Engines' Search Results When People Google for Suicide. *Health Communication* 32, 2 (Feb. 2017), 253–258. https://doi.org/10.1080/10410236.2015.1113484

[9] Mario Haim, Andreas Graefe, and Hans-Bernd Brosius. 2018. Burst of the Filter Bubble?: Effects of personalization on the diversity of *Google News*. *Digital Journalism* 6, 3 (March 2018), 330–343. https://doi.org/10.1080/21670811.2017.1338145

[10] Aniko Hannak, Piotr Sapiezynski, Arash Molavi Kakhki, Balachander Krishnamurthy, David Lazer, Alan Mislove, and Christo Wilson. 2013. Measuring personalization of web search. In *Proceedings of the 22nd international conference on World Wide Web - WWW '13*. ACM Press, Rio de Janeiro, Brazil, 527–538. https://doi.org/10.1145/2488388.2488435

[11] Jens Kr Steen Jacobsen and Ana María Munar. 2012. Tourist information search and destination choice in a digital age. *Tourism management perspectives* 1 (2012), 39–47.

[12] Rob Kitchin. 2017. Thinking critically about and researching algorithms. *Information, Communication & Society* 20, 1 (2017), 14–29.

[13] Chloe Kliman-Silver, Aniko Hannak, David Lazer, Christo Wilson, and Alan Mislove. 2015. Location, Location, Location: The Impact of Geolocation on Web Search Personalization. In *Proceedings of the 2015 ACM Conference on Internet Measurement Conference - IMC '15*. ACM Press, Tokyo, Japan, 121–127. https://doi.org/10.1145/2815675.2815714

[14] Yaniv Leviathan and Yossi Matias. 2018. Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone. http://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html

[15] lxml. 2017. lxml - Processing XML and HTML with Python. Retrieved April 18, 2019 from https://lxml.de/

[16] Judith Möller, Damian Trilling, Natali Helberger, and Bram van Es. 2018. Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society* 21, 7 (2018), 959–977.

[17] Jacob Ørmen. 2018. A discussion of research designs for assessing algorithmic curation. *The Routledge Handbook of Developments in Digital Journalism Studies* (2018).

[18] Bing Pan, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gay, and Laura Granka. 2007. In google we trust: Users' decisions on rank, position, and relevance. *Journal of computer-mediated communication* 12, 3 (2007), 801–823.

[19] Eli Pariser. 2011. *The Filter Bubble: What The Internet Is Hiding From You.* Penguin UK.

[20] Nick Seaver. 2014. Knowing algorithms.

[21] Dennis Sellers. 2018. Google Maps has 67% of the navigation app market compared to Apple Map's 11%. Retrieved April 19, 2019 from https://www.appleworld.today/blog/2018/7/10/google-maps-has-67-of-the-navigation-app-market-compared-to-apple-maps-11

[22] Joe Shaw and Mark Graham. 2017. An informational right to the city? Code, content, control, and the urbanization of information. *Antipode* 49, 4 (2017), 907–927.

[23] Cass R Sustein. 2009. Republic. com 2.0. *Nova Jersey: Princeton University Press* (2009).

[24] W3C. 2017. XML Path Language (XPath) 3.1. Retrieved April 18, 2019 from https://www.w3.org/TR/2017/REC-xpath-31-20170321/

[25] Michele Willson. 2017. Algorithms (and the) everyday. *Information, Communication & Society* 20, 1 (Jan. 2017), 137–150. https://doi.org/10.1080/1369118X.2016.1200645

[26] Zheng Xiang and Ulrike Gretzel. 2010. Role of social media in online travel information search. *Tourism Management* 31, 2 (April 2010), 179–188. https://doi.org/10.1016/j.tourman.2009.02.016

[27] Bo Yu and Guoray Cai. 2007. A query-aware document ranking method for geographic information retrieval. In *Proceedings of the 4th ACM workshop on Geographical information retrieval*. ACM, 49–54.