

# **The Lure of Google Trends' Objectivity: Knowledge Production through Data Assemblages**

A Critical Data Studies Approach towards Google Trends'  
non-neutrality, situatedness and epistemic ramifications

Charlotte Hannen (6562671)

Date: 13-06-2019

Supervisor: Alex Smit

Pre-Master New Media & Digital Culture 2018 – 2019 (4)

# The Lure of Google Trends' Objectivity: Knowledge Production through Data Assemblages

A Critical Data Studies Approach towards Google Trends'  
non-neutrality, situatedness and epistemic ramifications

## Abstract

As our contemporary society becomes increasingly datafied, digital tools to study societal phenomena through empirical data become more ubiquitous. A digital tool that is commonly used by scientists to study societal phenomena is Google Trends. Google Trends is an online service that records the index of search queries at a particular time within a chosen geographical area. Despite that there is an ongoing debate on Google Trends' objectivity, its research output is used as proxies for human behavior, public opinion and societal change. As the academic debate majorly focuses on uncertainties on whom Google Trends' data actually represents, this research aims to fill the gap between this debate by analyzing how the tool itself affects knowledge production through the paradigm of Critical Data Studies. Observing Google Trends through the framework of Critical Data Studies shows that this socio-technological system is conceived as a data assemblage whose interrelated processes are deeply intertwined. Additionally, this research enriches the framework of Critical Data Studies with the methodology of Tool Criticism, which shows that Google Trends' apparatuses *terms and comparison*, *filters*, *interest over time*, *interest by subregion*, *related topics* and *related queries* are conceivably non-neutral, non-transparent, situated and partially remain black boxes. Consequently, the tool's research output is considered as limited and biased, which affects how knowledge is produced through Google Trends. Subsequently, it is shown that the assemblage produces knowledge through a two-dimensional perspective in terms of both interface and inner-workings. Due to a fundamental power relation between Google and its users, Google itself primarily benefits from this perspective, while minority groups consequently suffer from this rather planar perspective. Additionally, it is shown that Google Trends is an exemplary socio-technological system whose epistemic relation with its users is augmented with an ontic relation. As Google Trends presupposed that its users can "explore what the world is searching," the tool implies that it serves as a virtual environment that reflects an online social reality. However, research has shown that this social reality is not deemed to be objective and neutral. In response to the non-neutrality and situatedness of Google Trends, the research finally calls for an inquiry that critically depicts on knowledge constitution through other socio-technological tools through the framework of Critical Data Studies.

**Keywords:** knowledge production – data assemblages – Critical Data Studies – Tool Criticism – Google Trends

# Table of contents

<b>1. Monitoring and forecasting societal phenomena</b>	<b>5</b>
1.1 The disputed objectivity of Google Trends	5
<b>2. Knowledge production through data assemblages</b>	<b>8</b>
2.1 Introducing the data assemblage	8
2.2 Inside the data assemblage	9
2.2.1 Black box characteristics and echo chambers	10
2.3 Critique on the data assemblage	10
2.3.1 The conventions that visualizations do by Kennedy et al.	12
2.3.2 The ambiguity of search queries	12
2.4 Epistemic ramifications of the data assemblage	13
2.4.1 An ontic relation between humans and cognitive artifacts	14
2.4.2 Harding's Standpoint Theory	14
<b>3. Utilizing methodologies of Critical Data Studies and Tool Criticism</b>	<b>15</b>
3.1 Explaining the corpus	15
3.2 Performing the methodologies	15
3.3 Limitations of the methodology and corpus	16
<b>4. Critically reflecting on the knowledge produced by Google Trends</b>	<b>17</b>
4.1 The engines of the data assemblage	17
4.2 Knowledge production through <i>terms and comparison</i>	18
4.3 Knowledge production through <i>filters</i>	19
4.4 Knowledge production through <i>interest over time</i> and <i>interest by subregion</i>	22
4.5 Knowledge production through <i>related topics</i> and <i>related queries</i>	23
4.6 Epistemic ramifications of Google Trends	24
<b>Conclusion</b>	<b>28</b>
<b>References</b>	<b>30</b>

## Appendices

Appendix 1. The looping effect	33
Appendix 2. Nine intellectual virtues	34
Appendix 3. ‘Engines of discoverability’ within Google Trends	35
Appendix 4. Example (anti-)programmes and neutral keywords	36
Appendix 5. List countries Google Trends in <i>filters</i>	37
Appendix 6. Google Searches 1999-2012	38
Appendix 7. Google Trends’ categories in <i>filters</i>	39
Appendix 8. Results term ‘rose’ when excluding categories	40
Appendix 9. Results term ‘rose’ when including categories	41
Appendix 10. Search Engine Market Share Worldwide May 2019	42
Appendix 11. Plagiarism awareness declaration form	43

## 1 Monitoring and forecasting societal phenomena

On May 1, 2019, Google announced the introduction of auto-delete controls.<sup>1</sup> This upcoming affordance allows users to automatically delete captured user activities and location histories and is scheduled to launch shortly.<sup>2</sup> As Google recently received bad press on its collection of user location histories and ambiguous privacy policies, the initiative is considered as the firm's response to these scandals.<sup>3,4,5</sup> However, since the affordance is not taken into operation yet, it is questionable how auto-delete controls affect Google's core business of selling targeted ad space. Moreover, enabling users to manage and delete personal data affects the output of additional Google services, such as Google Trends.<sup>6</sup>

In the past decade, the number of studies that utilize Google Trends to index societal phenomena has increased dramatically.<sup>7</sup> In "Ten years of research change using Google Trends", 657 of these studies are observed. The scientists Seung-Pyo Jun, Hyoung Sun Yoo and San Choi found that the focus of these studies has shifted from monitoring and surveillance towards forecasting purposes.<sup>8</sup> Moreover, it is shown that internet search data are used as proxies for human behavior, public opinion and societal change.<sup>9,10,11,12</sup> However, there is an ongoing debate on whether big data - such as Google Trends - are appropriate to index societal phenomena.

### 1.1 The disputed objectivity of Google Trends

Proponents of big data emphasize its advantages over e.g. survey data in terms of availability, costs and objectivity.<sup>13,14</sup> Google Trends' advantages over survey data regarding availability and costs are somewhat self-evident, as the tool's daily search query updates result in a continuous flow of free accessible data.<sup>15,16</sup> Despite that some academics may argue Google's affordability in terms of exploitation, the majority of academic debates focus on the tool's objectivity.

---

<sup>1</sup> David Monsees, and Marlo McGriff, "Introducing auto-delete controls for your Location History and activity data," The Keyword, accessed May 9, 2019, <https://blog.google/technology/safety-security/automatically-delete-data/>.

<sup>2</sup> Monsees, and McGriff, "Introducing auto-delete controls."

<sup>3</sup> "Google verzweeg bug waardoor gebruikersdata ingezien kon worden," Tech, NOS, accessed May 9, 2019, <https://nos.nl/artikel/2253973-google-verzweeg-bug-waardoor-gebruikersdata-ingezien-konden-worden.html>.

<sup>4</sup> "Klacht tegen Google vanwege 'ontfutselen' locatiegegevens," Tech, NOS, last modified November 11, 2018, <https://nos.nl/artikel/2261042-klacht-tegen-google-vanwege-ontfutselen-locatiegegevens.html>.

<sup>5</sup> "Mensen kunnen straks automatisch Google-data laten verwijderen," Tech, NOS, accessed May 9, 2019, <https://nos.nl/artikel/2282875-mensen-kunnen-straks-automatisch-google-data-laten-verwijderen.html>.

<sup>6</sup> "Explore," Google Trends, accessed May 2, 2019, <https://trends.google.com/trends/explore>.

<sup>7</sup> Seung-Pyo Jun, Hyoung Sun Yoo, and San Choi, "Ten years of research change using Google Trends: From the perspective of big data utilizations and applications," *Technological Forecasting and Social Change* 130 (2018): 69-71, <https://doi.org/10.1016/j.techfore.2017.11.009>.

<sup>8</sup> Jun, Yoo, and Choi, "Ten years of research change using Google Trends," 69.

<sup>9</sup> Jonathan Mellon, "Internet Search Data and Issue Saliency: The Properties of Google Trends as a Measure of Issue Saliency," *Journal of Elections, Public Opinion & Parties* 24, no. 1 (October 2013): 45, <https://doi.org/10.1080/17457289.2013.846346>.

<sup>10</sup> Jordan Wolf, "Trending in the Right Direction: Using Google Trends Search Data as a Measure of Public Opinion During a Presidential Election," (Thesis, Virginia State University, 2018).

<sup>11</sup> Jonathan Zhu, Xiaohua Wang, Jie Qin, Lingfei Wu, "Assessing Public Opinion Trends based on User Search Queries: Validity, Reliability, and Practicality," *The Annual Conference of the World Association for Public Opinion Research* (June 2012): 1-7.

<sup>12</sup> Ladislav Kristoufek, "Can Google Trends search queries contribute to risk diversification?," *Scientific reports* 3, no. 2713 (2013): 1.

<sup>13</sup> Mellon, "Internet Search Data and Issue Saliency," 45.

<sup>14</sup> Zhu, Qin and Wu, "Assessing Public Opinion Trends," 1-2.

<sup>15</sup> Google Trends, "Explore."

<sup>16</sup> Jun, Yoo, and Choi, "Ten years of research change using Google Trends," 84.

On one side of this debate, proponents consider Google Trends' objectivity as "relatively high" because the number of users is corresponding approximately to that of the population.<sup>17</sup> In addition, proponents prefer Google Trends over other research methods – in terms of objectivity – because of the users' lack of cognitive dissonance.<sup>18</sup> On the other side of the debate, critics question Google Trends' objectivity because it might solely reflect the opinion of the younger, more active and more educated section of the population.<sup>19,20</sup> In addition, critics question the users' lack of cognitive dissonance, because users may rapidly and emotionally response to gossip and negative news. As a result, critics claim that Google Trends' output reflects its users' irrational tendencies and is therefore biased.<sup>21</sup>

It can be concluded that the debate's overarching theme is how the research input might be distorted. More specifically, how the conclusions that are drawn from the research output are potentially biased by uncertainties on whom the data input represents. However, it is notable that, within the ongoing debate, there is limited knowledge on how Google Trends itself – as well as its translation, representation and visualization of the data – might bias its users' interpretations and assumptions of the data. In other words, the debate on Google Trends' objectivity lacks arguments on how the tool constitutes knowledge.

In order to complement the academic debate on Google Trends' objectivity, the research question posed here is: How does Google Trends produce knowledge as a data assemblage through the lens of Critical Data Studies and Tool Criticism? Which leads to the following sub-questions: How do Google Trends' apparatuses meet the elements of the theory on 'engines of discoverability'? How do Google Trends' apparatuses produce knowledge based on their ambiguity of keywords, visualization conventions and echo chambers? And what are the apparatuses' and assemblages' epistemic and ontic ramifications?

The aim of this research is to fill the gap between the academic debate concerning Google Trends' objectivity by critically analyzing how the tool constitutes knowledge. Therefore, this study particularly focuses on the interaction and relation between the researcher and Google Trends, through the lens of the humanities, by means of its interface, graphics and affordances. Accordingly, the research perspective and approach closely resemble to the paradigm that Dalton and Thatcher call Critical Data Studies.<sup>22,23</sup> This approach is considered as a research attitude that applies critical social theories to data, in order to depict how data are not plainly objective and neutral presentations of society, but are rather contingent and situated.<sup>24</sup> In this research, the social theories utilized for the purpose of Critical Data

---

<sup>17</sup> Jun, Yoo, and Choi, "Ten years of research change using Google Trends," 84.

<sup>18</sup> Jun, Yoo, and Choi, "Ten years of research change using Google Trends," 84.

<sup>19</sup> Zhu, Qin and Wu, "Assessing Public Opinion Trends," 1.

<sup>20</sup> Ernesto D'Avanzo, Giovanni Pilato, and Miltiadis Lytras, "Using twitter sentiment and emotions analysis of google trends for decisions making," *Program* 51, no. 3 (2017): 323-324.

<sup>21</sup> Jun, Yoo, and Choi, "Ten years of research change using Google Trends," 85.

<sup>22</sup> Craig Dalton, and Jim Thatcher, "Inflated granularity: The promise of big data and the need for a critical data studies," In *Presentation at the Annual Meeting of the Association of American Geographers, Tampa*, 2014.

<sup>23</sup> Craig Dalton, and Jim Thatcher, "What does a critical data studies look like, and why do we care? Seven points for a critical approach to 'big data'," *Society and Space* 29 (2014).

<sup>24</sup> Rob Kitchin, and Tracey Lauriault, "Towards critical data studies: Charting and unpacking data assemblages and their work," (2014): 5.

Studies are that of Tool Criticism by Karin Van Es, Maranke Wieringa and Mirko Schäfer and Hacking's theory on data assemblages.<sup>2526</sup> Second, this research aims to critically evaluate on Google Trends' epistemic ramifications through Richard Heersmink's approach on virtue epistemology, enriched by Sandra Harding's concept of Standpoint Theory.<sup>2728</sup> Finally, this study aims to raise awareness on how affordances such as auto-delete controls, shape the research output when indexing societal phenomena by means of Google Trends.

---

<sup>25</sup> Karin Van Es, Maranke Wieringa, and Mirko Tobias Schäfer, "Tool Criticism: From Digital Methods to Digital Methodology," In *Proceedings of the 2nd International Conference on Web Studies*, pp. 24-27, ACM, 2018.

<sup>26</sup> Kitchin, and Lauriault, "Towards critical data studies," 8-10.

<sup>27</sup> Richard Heersmink, "A virtue epistemology of the Internet: Search engines, intellectual virtues and education," *Social Epistemology* 32, no. 1 (2018): 1-12.

<sup>28</sup> Sandra Harding, *Objectivity and diversity: Another logic of scientific research* (University of Chicago Press, 2015).

## 2 Knowledge Production through data assemblages

### 2.1 Introducing the data assemblage

As mentioned in the introduction, this research positions itself within the paradigm of Critical Data Studies. As this approach facilitates as a framework to critically depict the situatedness, non-neutrality and subjectivity of data, it is considered as appropriate to reflect on the knowledge production by Google Trends through the lens of Critical Data Studies.<sup>29</sup> At the Association of American Geographers, Rob Kitchin called for the need for Critical Data Studies while pointing out the concept of ‘data assemblages’.<sup>30</sup> The geographer defines the term as, “a complex socio-technical system, composed of many apparatuses and elements that are thoroughly entwined, whose central concern is the production of a data.”<sup>31</sup> In addition to the system and infrastructure of the assemblage, it includes all “technological, political, social and economic apparatuses that frames their nature, operation and work.”<sup>32</sup> In order to chart and unpack a data assemblage, as well as discovering the apparatuses that might frame its nature, Rob Kitchin and Tracey Lauriault put forward theories of philosopher Ian Hacking.<sup>33</sup>

According to Kitchin and Lauriault, Hacking claims that there are two interrelated processes engaged within a data assemblage that are responsible for producing and legitimizing its data and corresponding apparatuses. In addition, these interrelated processes shape how the data assemblage performs in the world. Subsequently, this performance influences future repetition of the data, as well as the future constitution of the data assemblage. The interaction between the data and its representation thus leads to mutual changes. The first of the interrelated processes is that what Hacking terms ‘the looping effect’.<sup>34</sup> According to Kitchin and Lauriault, “the looping effect concerns how data are classified and organized: how a data ontology comes into existence and how it can reshape that which has been classified.”<sup>35</sup> The loop itself exists of five phases which are assigned to reform society through data ontology (as shown in appendix 1).<sup>36</sup> The second interrelated process is that of ‘engines of discoverability’. Hacking considers that these engines extend beyond simply methods. The figure below shows Hacking’s distinction between six engines of discovery and three derived engines (see figure 1).

---

<sup>29</sup> Kitchin, and Lauriault, "Towards critical data studies," 5.

<sup>30</sup> Kitchin, and Lauriault, "Towards critical data studies," 8.

<sup>31</sup> Kitchin, and Lauriault, "Towards critical data studies," 8.

<sup>32</sup> Kitchin, and Lauriault, "Towards critical data studies," 8.

<sup>33</sup> Kitchin, and Lauriault, "Towards critical data studies," 10.

<sup>34</sup> Kitchin, and Lauriault, "Towards critical data studies," 10.

<sup>35</sup> Kitchin, and Lauriault, "Towards critical data studies," 10.

<sup>36</sup> Kitchin, and Lauriault, "Towards critical data studies," 10.





Figure 1. Ian Hacking's theory on engines of discoverability<sup>3738</sup>

It appears that, collectively, these engines legitimize and reproduce the work the assemblage, as well as the assemblage itself.<sup>39</sup> However, before unpacking the assemblage through Hacking's theory, one should be aware of the assemblage's – and in this case Google Trends' – characteristics.

## 2.2 Inside the data assemblage

Google Trends is an online service that records the index of a particular search activity. Seth Stephens-Davidowitz and Hal Varian, who are both American economists, give a rather broad description of the tool saying, "the index measures the fraction of queries that include the term in question in the chosen geography at a particular time relative the total number of queries at that time."<sup>40</sup> Google Trends' inner-workings are generally described as a form of indexation that is based on the sum of *query share*. The share of a query is calculated as follows: the total query volume for a keyword within a chosen geographical area, dividend by the total number of queries during the selected time period within the chosen geographical area – with a maximum of 100.<sup>4142</sup> The query share formula explains that, when a certain query experiences a decrease in popularity through time, this indicated that there are fewer searches on this particular query as a percentage of all Google searches than there were previously.<sup>4344</sup>

The advantage of Google Trends is thus that the captured data indicates whether a search term is increasing in popularity over time within a chosen geographical area. Moreover, the tool allows its users to compare relative popularity.<sup>4546</sup> One of the disadvantages of Google Trends is its ambiguity of keywords (e.g. 'Apple' as a fruit type or a brand). However, research suggests that this ambiguity issue has a low impact on the query output, whenever the user selects an appropriate category (e.g. 'Food &

<sup>37</sup> Kitchin, and Lauriault, "Towards critical data studies," 10-12.

<sup>38</sup> Tracey Lauriault, "Data, Infrastructures and Geographical Imaginations," (PhD diss., Carleton University, 2012), 101.

<sup>39</sup> Kitchin, and Lauriault, "Towards critical data studies," 11.

<sup>40</sup> Stephens-Davidowitz, Seth, and Hal Varian, *A Hands-on Guide to Google Data* (Mountain View: Google Inc., 2014), 12, <http://people.ischool.berkeley.edu/~hal/Papers/2015/primer.pdf>

<sup>41</sup> Hyunyoung Choi, and Hal Varian, "Predicting the Present with Google Trends," *The Economic Record* 88 (June 2012): 3, <https://doi.org/10.1111/j.1475-4932.2012.00809.x>.

<sup>42</sup> Pedro Palos-Sanchez, and Marisol Correia, "The collaborative economy based analysis of demand: Study of Airbnb case in Spain and Portugal," *Journal of theoretical and applied electronic commerce research* 13, no. 3 (2018): 88.

<sup>43</sup> Since, in absolute numbers, the searches on every topic increase over time.

<sup>44</sup> Choi, and Varian, "Predicting the Present," 3.

<sup>45</sup> "Compare Trends search terms," Trends Help, Google, accessed March 18, 2019, <https://support.google.com/trends/answer/4359550?hl=en>.

<sup>46</sup> Stephens-Davidowitz, and Varian, *Hands-on Guide*, 13.

Drinks’ or ‘Business & Industrial’).<sup>47</sup> Another disadvantage are the tool’s black box characteristics, since its technical opacity and complexity hides and obfuscates its inner workings, which may shape the research output.<sup>4849</sup>

### 2.2.1 Black box characteristics and echo chambers

Daniel Golovin et al., who are researchers at Google, define a black box as any complex system that is easier to experiment with, rather than to actually understand.<sup>50</sup> An example of a Google affordance that is considered as a black box, is its ambiguous search infrastructure. As it enables users to unrestrictedly enter search queries, it is easy to experiment with. Nonetheless, users are unable to actually understand how Google ranks proposed sites. Is the engine biasing results that benefits its own commercial interests, or is it actually acting in “good faith” to serve its users?<sup>51</sup> In addition, Google’s search infrastructure is considered as a black box because it is personalizing results through machine-learning models.<sup>5253</sup> Consequently, different users automatically receive different results when entering an identical query. While some scientists argue that personalization of search queries increases exposure to differing perspectives, others claim that personalized search queries increase ideological segregation.<sup>5455</sup> Critics worry that this form of personalization might damage the democratic public sphere because, “like-minded individuals reinforce with their previously held beliefs.”<sup>56</sup> Cass Sunstein, who has been named a Harvard’s University Professor, refers to the concept of an individual being predominantly exposed to conforming opinions as ‘echo chambers’.<sup>57</sup> As this research is situated within the paradigm of Critical Data Studies, the concept of echo chambers is considered as a supporting resource to critically evaluate on Google’s situatedness and non-neutrality.

### 2.3 Critique on the data assemblage

As the characteristics of tools shape the research output – as in the case the characteristics of Google Trends – academics call for the critical inquiry of Tool Criticism.<sup>58</sup> The necessity of this inquiry derives from the datafication of our contemporary society. As our contemporary society becomes increasingly datafied, researchers have access to plenty of observational data. In order to collect, clean, process and visualize this data, digital methods tools are commonly utilized.<sup>59</sup> Van Es, Wieringa en Schäfer, who

---

<sup>47</sup> Stephens-Davidowitz, and Varian, *Hands-on Guide*, 15.

<sup>48</sup> Nicholas Diakopoulos, “Algorithmic Accountability Reporting: On The Investigation Of Black Boxes.” (PhD diss., Columbia Journalism School, 2013), 14.

<sup>49</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 26.

<sup>50</sup> Daniel Golovin et al., “Google vizier: A service for black-box optimization,” in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 1487, ACM, 2017.

<sup>51</sup> Frank Pasquale, *The black box society* (Harvard University Press, 2015), 9.

<sup>52</sup> Aniko Hannak et al., “Measuring personalization of web search,” in *Proceedings of the 22nd international conference on World Wide Web*, p. 527, ACM, 2013.

<sup>53</sup> Seth Flaxman, Sharad Goel, and Justin M. Rao, “Filter bubbles, echo chambers, and online news consumption,” *Public Opinion Quarterly* 80 (2016): 299.

<sup>54</sup> Hannak et al., “Measuring personalization of web search,” 527.

<sup>55</sup> Flaxman, Goel, and Rao, “Filter bubbles,” 298.

<sup>56</sup> Max Grömping, “‘Echo Chambers’ Partisan Facebook Groups during the 2014 Thai Election,” *Asia Pacific Media Educator* 24, no. 1 (2014): 53.

<sup>57</sup> Cass R. Sunstein, *Echo chambers: Bush v. Gore, impeachment, and beyond* (Princeton, NJ: Princeton University Press, 2001), 17.

<sup>58</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 24.

<sup>59</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 26.

are affiliated with the Utrecht Data School, question the non-neutrality of these tools.<sup>60</sup> In “Tool Criticism: From Digital Methods to Digital Methodology”, they emphasize the importance of critical reflexivity.<sup>61</sup> The academics plead for a reflexive and critical engagement with digital methods tools.<sup>62</sup> In their article, they connect disparate reflections on how tools affect research output under the label Tool Criticism, defining it as,

The critical inquiry of knowledge technologies used for research purposes. It reviews the qualities of the tool in light of the research activities and reflects on how the tool (i.e. its working mechanisms, anticipated use, interface and embedded assumptions) affects the research process and output.<sup>63</sup>

As acknowledged by Van Es, Wieringa and Schäfer, the concept of Tool Criticism is not considered to be a new phenomenon.<sup>64</sup> In fact, there are multiple similarities between the above mentioned definition of Tool Criticism and the definition of Digital Tool Criticism put forward by Koolen, Van Gorp and Van Ossenbruggen. The academics define Digital Tool Criticism as, “the reflection on the role of digital tools in the research methodology and the evaluation of the suitability of a given digital tool for a specific research goal.”<sup>65</sup> The aim of this concept is not to improve the tool’s performance, but to consider that its limitations impact the research goal.<sup>66</sup> Hence, Digital Tool Criticism focuses on the awareness of the impact of a tool on research methods, outcomes and interpretations.<sup>67</sup>

Despite various obvious similarities, there are two crucial differences between the definitions.<sup>68</sup> First, Van Es, Wieringa and Schäfer incorporate the importance of reflexivity on the interaction between the tool and the researcher. According to them, “considerations of, for instance, the tool’s complexity or intelligibility, or one’s lack of skills, often prove to be reasons for opting for particular tools.”<sup>69</sup> Therefore, they plead that the relation and interaction between the researcher and the tool affects how knowledge is produced and should therefore be reflected on, while Koolen, Van Gorp and Van Ossenbruggen seem to exclude the importance of this type of reflexivity.<sup>70</sup> Second, while the practice of Digital Tool Criticism does not aim to improve a tool’s performance, Van Es, Wieringa and Schäfer argue that Tool Criticism invites to “new and better tool development.”<sup>71,72</sup> As this research aims to

---

<sup>60</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 24.

<sup>61</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 25.

<sup>62</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 25.

<sup>63</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 26.

<sup>64</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 26.

<sup>65</sup> Marijn Koolen, Jasmijn van Gorp, and Jacco van Ossenbruggen, “Toward a model for digital tool criticism: Reflection as integrative practice,” *Digital Scholarship in the Humanities* 34, no. 2 (2018): 381.

<sup>66</sup> Koolen, Van Gorp, and Van Ossenbruggen, “Toward a model for digital tool criticism,” 381-382.

<sup>67</sup> Koolen, Van Gorp, and Van Ossenbruggen, “Toward a model for digital tool criticism,” 382.

<sup>68</sup> In the context of this research.

<sup>69</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 26.

<sup>70</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 26.

<sup>71</sup> Van Es, Wieringa, and Schäfer, “Tool Criticism,” 26.

<sup>72</sup> Koolen, Van Gorp, and Van Ossenbruggen, “Toward a model for digital tool criticism,” 381.

analyze how knowledge is produced through Google Trends, by analyzing the interaction between the user and the tool, this research builds upon the foundation of Tool Criticism as presented by Van Es, Wieringa en Schäfer.

### **2.3.1 The conventions that visualizations do by Kennedy et al.**

As Google Trends interacts with its users through data visualizations, this research enriches the concept of Tool Criticism with the theory of “The work that visualisation conventions do” by Helen Kennedy et al.<sup>73</sup> In this article, the sociologists point out the persuasiveness of data visualizations and criticize their ability to illustrate “our complex social world,” as they depict four common visualization conventions that create a sense of transparency, facticity and objectivity.<sup>74,75</sup> First, the sociologists criticize visualizations that contain two-dimensional viewpoints. The two-dimensional properties disguise perspectives, and therefore create the illusion of objectivity, which leaves the viewer under the false impression that the visualization provides the overall picture.<sup>76</sup> Second, the sociologists found that plain geometric shapes are commonly used in data visualizations. The advantage of visualizations with a simple pattern is that it yields the feeling of order, which enables the viewer to easily make sense of the visualization.<sup>77</sup> However, the drawback is that the data is dramatically and systematically simplified, because only some characteristics of the data are highlighted.<sup>78</sup> The same drawback applies to the third conventions of the use of clean lay-outs. Clean layouts create an aura of simplicity and therefore disguise the data’s complexity. In the process of cleaning the lay-out, the data is – yet again – dramatically simplified.<sup>79</sup> Finally, Kennedy et al. point out that including data sources in the visualization, creates a sense of objectivity and transparency and is therefore considered to be persuasive. As data sources suggest that the engineer has been truthful to the data, users may not feel to urge to check the data sources.<sup>80</sup> By critically evaluating on the ideological work that these four visualization conventions do as described by Kennedy et al., it becomes clear how Google Trends’ visualizations persuade and interact with its users.

### **2.3.2 The ambiguity of search queries**

Furthermore, Google Trends’ users interact with the tool through search terms. Hence, theories concerning queries and keywords design are queried. According to sociologists Latour and Akrich, one with social research purposes should formulate a search query in which both programmes and anti-programmes are deployed. Programmes refer to efforts made for the purpose of promoting and putting

---

<sup>73</sup> Helen Kennedy et al., “The work that visualisation conventions do,” *Information, Communication & Society* 19, no. 6 (2016): 715-735.

<sup>74</sup> Kennedy et al., “The work,” 723.

<sup>75</sup> Kennedy et al., “The work,” 715.

<sup>76</sup> Kennedy et al., “The work,” 723-724.

<sup>77</sup> Kennedy et al., “The work,” 724-725.

<sup>78</sup> Kennedy et al., “The work,” 727.

<sup>79</sup> Kennedy et al., “The work,” 729.

<sup>80</sup> Kennedy et al., “The work,” 731.

forward a particular campaign, project or proposal. In contrast, anti-programmes oppose these projects or efforts through keywords.<sup>8182</sup> Richard Rogers, who holds the Chair in New Media & Digital Culture at the University of Amsterdam, pleads for a third type of keywords, namely projects or efforts made at being neutral.<sup>83</sup> Neutral keywords are distinct undertakings made not to side with a programme or anti-programme (e.g. newspapers that advice its reporters to employ or avoid a particular language).<sup>84</sup> The categorization of the research output as a programme, anti-programme or neutral (combination of) keyword(s) may facilitate to draw conclusions on the output's epistemic ramifications.

## 2.4 Epistemic ramifications of the data assemblage

Historically, the interaction and relation between humans and computers is primarily considered as epistemic.<sup>8586</sup> Various scientists claim that the computer's ability to solve problems and process information, extends our human cognition.<sup>8788</sup> In "The Epistemology and Ontology of Human-Computer Interaction", Philip Brey compares the human cognitive abilities of memory, interpretation, search and conceptual thought to those of computers.<sup>89</sup> Subsequently, the philosopher concludes that computers are able to perform cognitive tasks autonomously and therefore extend human cognitive ability saying,

The functional relation that computers, as cognitive artifacts, have to their human users is hence that they extend cognition. Specifically, they extend the memory, interpretation, search, pattern matching and higher-order cognitive abilities of human beings.<sup>90</sup>

However, humans should be attentive to the potential epistemic risks for information-seeking and knowledge acquisition when utilizing cognitive artifacts, particularly in the case of internet-based sources.<sup>91</sup> Richard Heersmink, who has been awarded with a PhD Philosophy of Cognitive Science, supports this claim with the argument that personalized rankings – as mentioned earlier in this theoretical framework – “may nudge one towards a wrong path of enquiry.”<sup>92</sup> Given the epistemic drawbacks of online search engines, Heersmink calls for the ‘virtue epistemology approach’. This approach focuses

---

<sup>81</sup> Madeleine Akrich, and Bruno Latour, "The De-Description of Technical Objects," in *Shaping Technology / Building Society: Studies in Sociotechnical Change*, ed. Wiebe Bijker & John Law (Cambridge, MA: MIT Press, 1992), 205-224.

<sup>82</sup> Richard Rogers, "Foundations of Digital Methods: Query Design," in *The Datafied Society: Studying Culture Through Data*, eds. Mirko Tobias Schäfer and Karin van Es, 75-94, (Amsterdam: Amsterdam University Press, 2017), 82-87.

<sup>83</sup> Universiteit van Amsterdam, "dhr. prof. dr. R.A. (Richard) Rogers," accessed March 24, 2019, <http://www.uva.nl/profiel/t/o/r.a.rogers/r.a.rogers.html?1553443885418>.

<sup>84</sup> Rogers, "Foundations of Digital Methods," 82-87.

<sup>85</sup> Philip Brey, "The epistemology and ontology of human-computer interaction," *Minds and Machines* 15, no. 3-4 (2005): 383.

<sup>86</sup> Heersmink, "A virtue epistemology of the Internet," 1.

<sup>87</sup> Brey, "The epistemology and ontology," 383.

<sup>88</sup> Marshall McLuhan, *Understanding Media: The Extensions of Man* (New York: McGraw-Hill, 1966).

<sup>89</sup> Brey, "The epistemology and ontology," 384-390.

<sup>90</sup> Brey, "The epistemology and ontology," 390.

<sup>91</sup> Heersmink, "A virtue epistemology of the Internet," 1.

<sup>92</sup> Heersmink, "A virtue epistemology of the Internet," 6.

on the agent's cognitive agency and character.<sup>93</sup> In his article "A Virtue Epistemology of the Internet: Search Engines, Intellectual Virtues and Education", Heersmink depicts nine intellectual virtues, with differing cognitive characters, that minimize epistemic risks (see appendix 2).<sup>94</sup> When interacting with a cognitive artifact, people should adopt the attitude of one of the intellectual virtues to limit epistemic ramifications.<sup>95</sup> As Heersmink focuses on the agent's cognitive agency and characters, the approach deviates from the traditional analytic epistemology because it lacks focus on epistemic justification and the nature of truth.<sup>96</sup>

#### **2.4.1 An ontic relation between humans and cognitive artifacts**

In addition, it is worth mentioning that it is recently argued that the epistemic relation between humans and cognitive artifacts has been augmented with an ontic relation.<sup>97</sup> Brey claims that, as computers generate virtual interactive environments that offer new structures to experience, they are considered as "an augmentation of the world as it existed before."<sup>98</sup> Obviously, these are not physically real structures. However, Brey argues that the virtual interactive environment is meaningful to humans, "sometimes as much as their physical equivalents."<sup>99</sup> This is demonstrated by the social structures that emerge through networked computers. In this case, the internet provides an online social reality.<sup>100</sup>

#### **2.4.2 Harding's Standpoint Theory**

Finally, the theories mentioned are enriched by Sandra Harding's critical approach of Standpoint Theory to depict Google Trends' epistemic ramifications. In her book *Objectivity and Diversity: Another logic of scientific research*, the feminist claims that "knowledge is always socially situated."<sup>101</sup> In other words, whomever may be concerned with knowledge production procedures should be aware of how the used perspective implicates knowledge. Moreover, the parties concerned should realize that some may benefit from the perspective in which knowledge is produced, whilst others do not. The inequality of who do – and do not – benefit from the perspective in which knowledge is produced reflects the power relation.<sup>102</sup> Harding's theory provides a window into how epistemic ramifications depend on how the used perspective implicates knowledge. Therefore, this theory is considered as the lens through which Google's perspectives – in which knowledge is produced and epistemic ramifications emerge – are criticized.

---

<sup>93</sup> Heersmink, "A virtue epistemology of the Internet," 2.

<sup>94</sup> Heersmink, "A virtue epistemology of the Internet," 3-4.

<sup>95</sup> Heersmink, "A virtue epistemology of the Internet," 3.

<sup>96</sup> Heersmink, "A virtue epistemology of the Internet," 1.

<sup>97</sup> Brey, "The epistemology and ontology," 383.

<sup>98</sup> Brey, "The epistemology and ontology," 395.

<sup>99</sup> Brey, "The epistemology and ontology," 395.

<sup>100</sup> Brey, "The epistemology and ontology," 396.

<sup>101</sup> Harding, *Objectivity and diversity*, " 150.

<sup>102</sup> Harding, *Objectivity and diversity*, "

### 3 Utilizing methodologies of Critical Data Studies and Tool Criticism

#### 3.1 Explaining the corpus

As this research depicts how Google Trends produces knowledge through the lens of Critical Data Studies and Tool Criticism, the tool's intertwined apparatuses, whose concern it is to produce data, are firstly analyzed. The apparatuses that are concerned with the production of data, and are hence considered as the majority of the corpus of this research are, 1) *interest over time*: the apparatus that counts the volume of phenomena, which is subsequently visualized in a line graph, 2) *interest by subregion*: counts the volume of a phenomena within a particular geographical area, which is subsequently visualized in a map, 3) *related topics*: the apparatus counts the number of related topics based on the user's entered term, which are subsequently categorized as 'top' or 'rising' and are finally indexed from 0 – 100 or translated to percentages, and finally 4) *related queries*: the apparatus counts the number of related queries based on the user's entered term, which are subsequently categorized as 'top' or 'rising' and are finally indexed from 0 – 100 or translated to percentages. Collectively, these apparatuses produce data in the overarching interest over Google Trends.

It is notable that these four apparatuses jointly come into force as a result of the input entered by Google Trends' users through the apparatuses *terms and comparison* and *filters*, which are thus also considered as the corpus of this research. Google Trends affords its users to enter – a combination of – verbs, articles, nouns, adjectives, numerals, prepositions, adverbs, conjunctions, pronouns and interjection through the apparatus *terms and comparison*. Moreover, Google Trends presents proposed terms and comparisons on its homepage. However, as this research focuses on the interaction between the users and the tool, it is presumed that users enter a search query themselves. Therefore, Google Trends' proposed terms and comparison as presented on its homepage are excluded from this research. Additionally, users are offered the possibility to specify the research output through *filters* in terms of countries, period of time, categories and type of search.

#### 3.2 Performing the methodologies

First, the methodology of Critical Data Studies is assessed on the apparatuses *interest over time*, *interest by subregion*, *related topics* and *related queries* through Hacking's theory on 'engines of discoverability', as these engines are concerned with the production of data. All four apparatuses' inner-workings are outlined through Hacking's framework as shown in a table in appendix 3.<sup>103104</sup> As for the analysis, the four apparatuses' connections and inner-workings are summarized based on this table.

---

<sup>103</sup> Kitchin, and Lauriault, "Towards critical data studies," 10-12.

<sup>104</sup> Lauriault, "Data Infrastructures and Geographical Imaginations," 101.

However, as the majority of Google Trends' underlying technical system is considered as a black box, the extend of thorough research through Hacking's framework presumably differs per apparatus.

Subsequently, theories that contribute to the methodology of Tool Criticism are applied to the apparatuses *terms and comparison*, *filters*, *interest over time*, *interest by subregion*, *related topics* and *related queries*. First, to critically reflect on the apparatus *search terms and comparison*, theories on the ambiguity of keywords are applied to the mechanisms' inner workings. Second, to reflect on the *filters*, its default options – as well as the options itself – are critically analyzed in terms of objectivity, neutrality and situatedness. Subsequently, to reflect on the apparatuses *interest over time* and *interest by subregion*, the conventions of the displayed visualizations are described through the theory of Kennedy et al.<sup>105</sup> Furthermore, this research critically reflects on the fifth and sixth apparatuses *related topics* and *related queries*, through the theories on echo chambers and black box characteristics as outlined in the theoretical framework.

In addition, the epistemic ramification of the inner-workings and interfaces of *terms and comparison*, *filters*, *interest over time*, *interest by subregion*, *related topics* and *related queries* – as well as the assemblage itself – are explained through a combination of Heersmink's 'virtue epistemology approach', Brey's theory on cognitive artifacts and Harding's Standpoint Theory.

### **3.2 Limitations of the methodology and corpus**

As a consequence of the choices made on whether or not to include certain corpus and methodologies, this research has several limitations. In terms of the corpus, it is considered a limitation that the majority of Google Trends' underlying technical system is considered as a black box. Although this might contribute to arguments on how the tool's lack of transparency affects how knowledge is produced, this limits the scope of the research. In terms of methodology, the method of Critical Data Studies is considered to fall short in terms of insights into Google Trends' users. As a consequence, this research can only assume how users might (mis-)interpret information and produce limited and biased knowledge through the tool.

---

<sup>105</sup> Kennedy et al., "The work," 715-735.



– analysis and results

## 4 Critically reflecting on the knowledge produced by Google Trends

This chapter critically reflects on knowledge production through Google Trends by outlining the assemblage's mechanisms, questioning the tool's objectivity, transparency and neutrality and describing its epistemic and ontic ramifications.

### 4.1 The engines of the data assemblage

As this research considers Google Trends as a data assemblage, the research implicates that the tool's apparatuses and elements are deeply intertwined. In order to depict the intensity of the connection between these elements, as well as elaborating on the apparatuses' individual systems, those whose concern it is to produce data are outlined through Hacking's theory of data assemblages (for table, see appendix 3).<sup>106107</sup>

It was found that the systems of *interest over time* and *interest by subregion* show several similarities in terms of data processing. Both systems are designed in a way that they count the volume of entered phenomena and subsequently turn those counts into an index from 0 – 100. As a result of the indexation, the working mechanisms of *interest over time* and *interest by subregion* are able to define the relationships between measurements by means of relative popularity. The same applies for the apparatuses *related topics* and *related queries*, as they both count volumes of phenomena and subsequently classify these counts before converting them into indexes and percentages. The working mechanisms of *related topics* and *related queries* define relationships between measurements by means of relative popularity and percentages fluctuations.

Moreover, there are similarities between the elements *interest over time*, *interest by subregion*, *related topics* and *related queries* as their inner-workings partially remain black boxes. Due to the lack of data sources and an overview of the utilized methodologies, it is uncertain which (scientific) knowledge is applied to the algorithms that process and visualize the data of these four apparatuses. Another similarity is that it is assumed that all four apparatuses are able to identify that what might be expected, as Google generally receives tremendous amounts of data on a daily basis – and has been receiving for several years.<sup>108</sup> Based on this captured data, predictive algorithms might map future expectations and thus forecast e.g. how shifting seasons change the content of entered terms and how this might affect future results of *interest over time*, *interest by subregion*, *related topics* and *related queries*. Additionally, the four mechanisms show similarities in terms anticipating on norms created by e.g. the media. As soon as media devote attention on a particular topic – and thus create a norm on the importance of this topic – Google might fit this norm by proposing completions to its users surrounding the particular topic. Due to the proposed completions, the number of Google searches on the topic might

---

<sup>106</sup> Kitchin, and Lauriault, "Towards critical data studies," 10-12.

<sup>107</sup> Lauriault, "Data Infrastructures and Geographical Imaginations," 101.

<sup>108</sup> "Google Search Statistics," Trends and More, Internet Live Stats, accessed June 7, 2019, <https://www.internetlivestats.com/google-search-statistics/>.

reflect the importance determined by the media within a certain period of time or within a particular geographical area. Another similarity is that there might be possibilities for minority groups to take over control of the data displayed by *interest over time*, *interest over subregion*, *related topics* and *related queries*. However, due to the total amount of Google searches, it is considered virtually impossible for minority groups to highlight their beliefs through Google Trends.

Furthermore, it is notable that Google Trends’ components of *interest over time*, *interest by subregion*, *related topics* and *related queries* collectively reproduce the assemblage itself as they automatically process data in the overarching interest of Google Trends. Finally, it is striking that all four apparatuses jointly come into force as a result of the input entered by Google Trends’ users through *terms and comparison* and *filters*.

### 4.2 Knowledge production through *terms and comparison*

As mentioned in the theoretical framework, it is argued that users may emotionally and rapidly response to gossip and negative news.<sup>109</sup> It is thus conceivable that whenever someone aims to analyze the discourse surrounding e.g. women in The Netherlands, one may simply enter the term ‘vrouwen’ (Dutch term for women, translated by author), since this is the term that both feminists and the media use when referring to this type of gender. In this case, the term ‘vrouwen’ thus serves as both a programme and neutral keyword.<sup>110</sup> As a result, the information the user receives for analyzing the discourse surrounding women might be framed as both positive and neutral (see figure 2).

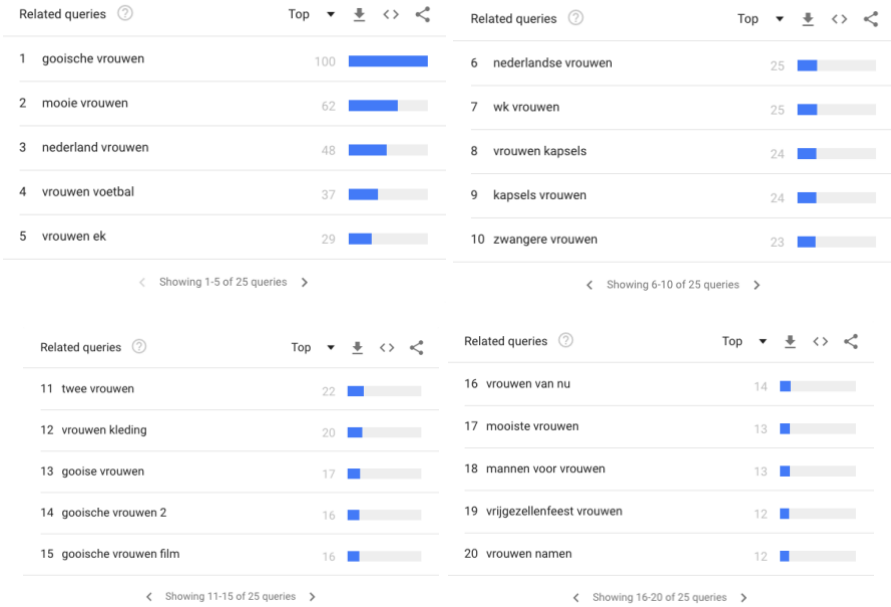


Figure 2. Results of *related queries* when entering programme and neutral keyword ‘vrouwen’ <sup>111112</sup>

<sup>109</sup> Jun, Yoo, and Choi, “Ten years of research change using Google Trends,” 85.  
<sup>110</sup> As feminists promote ‘vrouwen’ and the media does not side with the topic of ‘vrouwen’.  
<sup>111</sup> Keyword entered: “vrouwen”. Filters: “Netherlands”, “2004-present”, “All categories” and “Web search”. Related queries: “Top”.  
<sup>112</sup> Images acquired from <https://trends.google.com/trends/explore?date=all&geo=NL&q=vrouwen>

The user's discourse changes when including knowledge derived from anti-programmes concerning the subject of women, such as the term 'wijven' (Dutch term for women that is considered as disparaging, translated by author, see figure 3).

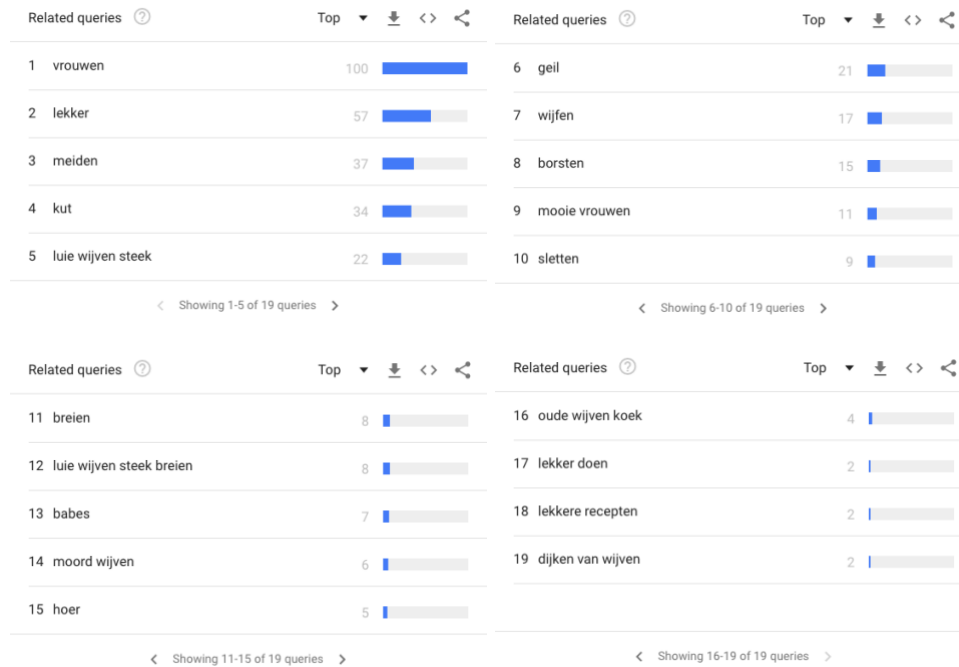


Figure 3. Results of *related queries* when entering anti-programme 'wijven' <sup>113114</sup>

One should thus acknowledge that including anti-programmes, that might be considered as controversial, expand the discourse surrounding certain topics and thus may enrich the knowledge produced (for example of African American connotations entered in Google Trends, see appendix 4).

As Google Trends allows its users to only enter one or two terms, users might solely search for either programmes, anti-programmes or neutral keywords. As a result, users might subconsciously absorb research output that either promotes, opposes or does not side with the entered query. In that case, the knowledge produced is considered as limited because it does not include the overall picture or entire discourse of a certain topic.

### 4.3 Knowledge production through *filters*

After entering one or two keywords, users are offered the possibility to specify the research output through several filters. First, one is offered the possibility to specify research input in terms of countries. The user can choose between 250 countries or adopt the worldwide filter (see appendix 5). It is noteworthy that Google Trends presents more countries than there are internationally recognized as autonomous states. This is due to that Google Trends includes countries that are not internationally

<sup>113</sup> Keyword entered: "wijven". Filters: "Netherlands", "2004-present", "All categories" and "Web search". Related queries: "Top".

<sup>114</sup> Images acquired from <https://trends.google.com/trends/explore?date=all&geo=NL&q=wijven>

recognized as autonomous states, such as Taiwan (see appendix 5).<sup>115</sup> Users who are aware of Taiwan's transmission to the People's Republic of China in 1945, might choose to filter China assuming that the research output then includes results on former Taiwan as well.<sup>116</sup> Therefore, Google Trends' affordance of filtering unrecognized autonomous states, may affect the research output and thus affect how knowledge is constituted.

Additionally, users are offered the possibility to filter the research output through various periods of time, as shown in figure 4.

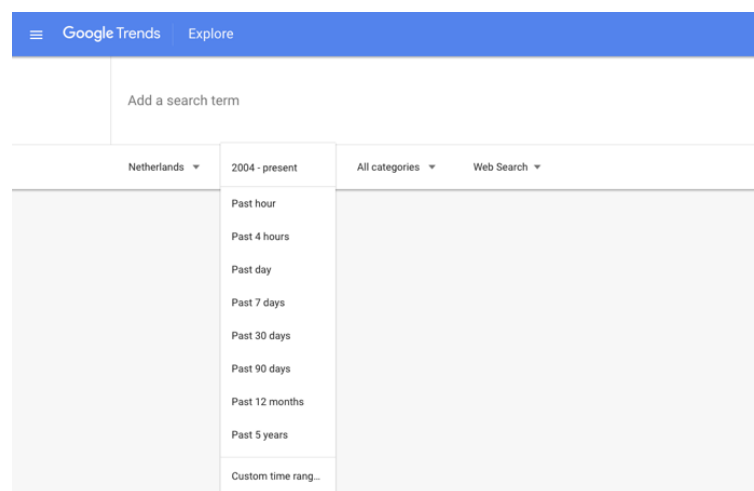


Figure 4. Filtering possible periods of time <sup>117</sup>

One who aims to research the popularity of a query should however realize that the number of Google searches in general has grown explosively from rounded off 86 billion searches in 2004 to rounded off 1,22 trillion searches in 2012 (see appendix 6).<sup>118</sup> As there were less Google searches in 2004 in general, this might negatively affect the sum of query share of the scaled topics in 2004, considering that the formula includes the total number of Google searches from 2004 – present. To put into other words, one could assume that if the total number of searches in 2004 and 2012 are divided over approximately the same quantity of topics, that the sum of query share might be generally higher in 2012 due to the increase of the total Google searches over time. When employing this filter, users should thus consider how the increase of total Google searches might limit the sum of query share and how this might affect how knowledge is produced.

<sup>115</sup> "Taiwan," Nederland wereldwijd, Koninkrijk der Nederlanden, accessed June 7, 2019, <https://www.nederlandwereldwijd.nl/landen/taiwan>.

<sup>116</sup> Agence France-Presse, "Chronology of Taiwan-China Relations Since 1945," *New York Times*, March 18, 2000, <https://archive.nytimes.com/www.nytimes.com/library/world/asia/031800china-taiwan-chron.html>.

<sup>117</sup> Image acquired from <https://trends.google.com/trends/explore?date=all&geo=NL>.

<sup>118</sup> Internet Live Stats, "Google Search Statistics."

Furthermore, Google Trends allows users to specify research output by selecting categories (see appendix 7). As mentioned in the theoretical framework, these categories are supposed to eliminate the ambiguity of keywords and are particularly useful when entering homonyms, such as the term ‘rose’. When entering the term without filtering any categories, the tool seems to combine results on – among others – the flower called rose, the past tense of the verb ‘rise’, the color rose and the last name of actress Ruby Rose (see appendix 8). Whenever filtering the research output on e.g. Home & Garden and People & Society, the homonyms are majorly eliminated (see appendix 9). It can be concluded from this that the tool’s research output might be biased when users enter a homonym without specifying their input with the appropriate category. Consequently, users might absorb ambiguous results, which might affect the way in which knowledge is produced.

Finally, Google Trends offers its users the possibility to specify research output per Google service in which terms are entered, as shown in figure 5.

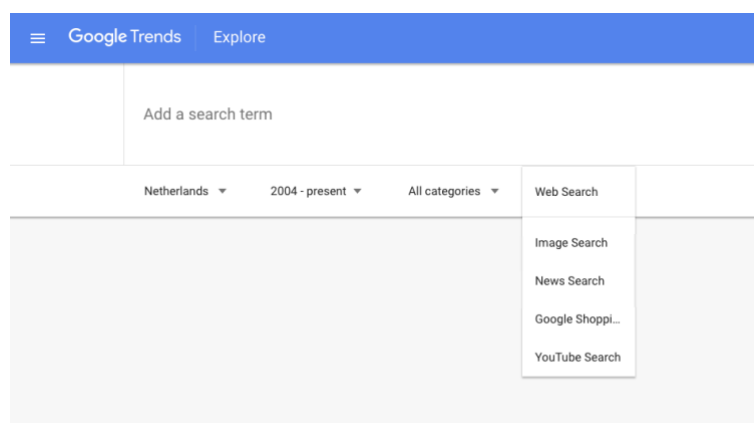


Figure 5. Filtering possible Google services <sup>119</sup>

It is first notable that the service Google Trends – which is also considered as a search engine – is not included in the filter. Consequently, Google Trends prevents its users from analyzing what trends are frequently searched for. In addition, it is arguable whether the term Web Search might be misleading as it may imply that this filter shows what is searched for on the entire web. Although Google had a search engine market share of 92% worldwide in May 2019, Google Trends’ research output does not include the remaining 8% of all web search (see appendix 10) .<sup>120</sup> Therefore, users of Google Trends should realize that the tool does not allow its users to “explore what the world is searching for,” but only allows its users to analyze what the majority of the world is searching for through the Google services as shown in figure 5.<sup>121</sup>

<sup>119</sup> Image acquired from <https://trends.google.com/trends/explore?date=all&geo=NL>.

<sup>120</sup> “Search Engine Market Share Worldwide,” StatCounter, accessed June 8, 2019, <http://gs.statcounter.com/search-engine-market-share#monthly-201905-201905-bar>.

<sup>121</sup> Google Trends, “Explore.”

#### 4.4 Knowledge production through *interest over time* and *interest by subregion*

As the apparatuses *interest over time* and *interest by subregion* interact with users through data visualizations, the theory of visualization conventions by Kennedy et al. is applied to critically reflect on the knowledge produced by these apparatuses through the lens of Tool Criticism.<sup>122</sup> The following table describes whether the visualizations, as presented by *interest over time* and *interest by subregion*, contain any conventions that might create a sense of transparency, facticity and objectivity (see table 1).

Apparatus	Two-dimensional Creating a sense of objectivity	Simple geometric shapes Creating a sense of facticity	Clean lay-outs Creating a sense of facticity	Data sources Creating a sense of transparency
<b>Interest over time</b>	Line graph contains two-dimensional viewpoints	The line graph is considered as a simple geometric shape It enables users to easily make sense of the graph (growth or decrease over time)	Clean layout, simple axes and scales: x-axis from max. 1/1/2014 to present y-axis from 0 to 100	Data source not available Note on 1/1/11 and 1/1/16 that data collection system improved
<b>Interest by subregion</b>	Interactive map which contains two-dimensional viewpoints	Maps are considered as visualizations that have a simple geometric shape It enables users to make sense of the graph since the intensity of the color indicates the relative score of a subregion	Clean lay-out Ranking list of subregions that supports the scores of the interactive map No axes and scales because it's a map	Data source not available

Table 1. Overview visualization conventions in *interest over time* and *interest by subregion* <sup>123</sup>

It can be inferred from the table that the visualizations presented by the apparatuses *interest over time* and *interest by subregion* meet three out of the four conventions as discussed by Kennedy et al.<sup>124</sup> Consequently, the visualizations displayed by *interest over time* and *interest by subregion* might create a sense of objectivity, as their two-dimensional viewpoints give its users the impression of what feminist Donna Haraway calls the “god-trick of seeing everything from nowhere.”<sup>125</sup> The two-dimensional viewpoint thus disguises any other perspectives, which might leave Google Trends’ users under the false impression that what is shown in the graph and map, is the whole truth. Moreover, the graph and map displayed by *interest over time* and *interest by subregion* create a sense of facticity, as their simple

<sup>122</sup> Kennedy et al., "The work," 715-735.

<sup>123</sup> Kennedy et al., "The work," 715-735.

<sup>124</sup> Kennedy et al., "The work," 715-735.

<sup>125</sup> Donna Haraway, "Situated knowledges: The science question in feminism and the privilege of partial perspective," *Feminist studies* 14, no. 3 (1988): 581.

geometric shapes might leave users under the impression that they ‘see’ the data directly, in their pure and clean form.<sup>126</sup>

Of greater concern is that both *interest over time* and *interest by subregion*, rather than creating a sense of transparency, are non-transparent by any means as they do not include or refer to any data sources. Although the apparatus *interest over time* refers to a certain data collection, when noting that improvements were applied to its data collection system on 1/1/11 and 1/1/6, there are uncertainties on what exact improvement were applied to what exact data collection system. Moreover, one can only wonder if these improvement were also applied to data captured before the dates of 1/1/11 and 1/1/16. If that is not the case, data before and after the dates of 1/1/11 and 1/1/16 might be inappropriate for comparing and contrasting purposes.

#### **4.5 Knowledge production through *related topics* and *related queries***

Google Trends shows individuals what topics and queries were searched for by other users who have initially searched for the same term as entered by this individual, through the apparatuses *related topics* and *related queries*. The apparatuses’ systems are designed in a way that they count the volume of chosen topics and queries that were proposed to Google Trends’ users through Google Autocomplete.

Google Autocomplete predicts searches based on what the user is typing, the user’s previous related search queries and the popular topics within the user’s geographical area.<sup>127</sup> As these proposed completions entail previous related searches, users are exposed to content conforming their personal opinion. Moreover, since Google Autocomplete channels its users towards popular topics within the same geographical area, it can be argued that an ideological segregation appears on regional level.<sup>128</sup> It can be inferred from this that Google Autocomplete’s characteristics correspond with the characteristics of an echo chamber as outlined in the theoretical framework. Therefore, it may be concluded that Google Autocomplete’s inner-workings constitute an echo chamber.

As Google Trends proposes completions to its users, and the way in which users respond to these suggestions are registered by *related topics* and *related queries*, these apparatuses naturally process registrations that are affected by users’ echo chambers. In fact, one could question Google Trends’ neutrality, as the proposed completions channels its users towards certain topics rather than assuming an impartial role. To put it even more strongly, one might criticize Google Trends’ situatedness as Google channels it users towards topics and queries that one is supposed to search for, presumably to fit norms (see figure 6).

---

<sup>126</sup> Kennedy et al., "The work," 729.

<sup>127</sup> "Search using autocomplete," Google Search Help, Google, accessed October 22, 2018, <https://support.google.com/websearch/answer/106230?co=GENIE.Platform%3DAndroid&hl=en>.

<sup>128</sup> Flaxman, Goel, and Rao, "Filter bubbles," 298.

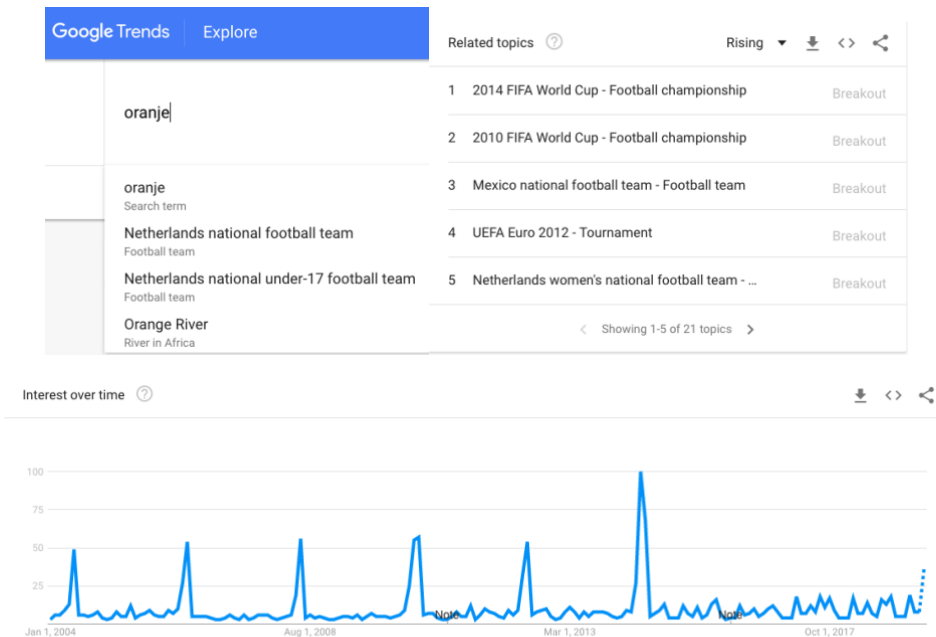


Figure 6. Proposed completions to fit norms <sup>129</sup>130

As the 2019 FIFA Women’s World Cup starts in June, 2019, Google Trends proposes completions when entering the term ‘oranje’ (Dutch term for orange which signifies national pride, translated by author) that refer to soccer rather than simply the color orange. The dotted line presented by *interest over time*, contains partially incomplete data but is expected to rise in June 2019. As the 2019 FIFA Women’s World Cup kicks off within the same month, a norm is created by means of this event. It is conceivable that Google Trends proposes soccer-related topics and queries to fit this norm. Consequently, users are channeled towards certain norms that might affect the way in which knowledge is constituted.

#### 4.6 Epistemic ramifications of Google Trends

To complement the foregoing paragraphs that critically reflect on knowledge constitution through Google Trends, it is now argued how the perspective used by the tool emerges epistemic ramifications and who does – and does not – benefit from this perspective.

From what has been explained in the foregoing paragraphs, it may be concluded that the perspective used by Google Trends can be labeled as two-dimensional in terms of both interface and inner-workings. In terms of interface, Google Trends uses this perspective in its graphs, maps and qualitative results.<sup>131</sup> As the interface is simply two-dimensional, it disguises any other perspectives. The drawback of this perspective is that it thus creates the lure of objectivity.<sup>132</sup> Users might benefit

<sup>129</sup> Keyword entered: “oranje”. Filters: “Netherlands”, “2004-present”, “All categories” and “Web search”.

<sup>130</sup> Images acquired from <https://trends.google.com/trends/explore?date=all&geo=NL&q=oranje>

<sup>131</sup> Qualitative results of words in *related topics* and *related queries* are considered as two-dimensional as well.

<sup>132</sup> Haraway, “Situated knowledges,” 581.



from this perspective in a way that they can easily make sense of the tool’s graphs, maps and qualitative results. However, users who do not critically reflect on the tool’s interface, might draw limited conclusions as they interpret the information given by Google Trends as if it were the whole truth.

In terms of inner-workings, the foregoing paragraphs have shown that Google Trends’ perspective might be considered as two-dimensional as well, as the tool’s neutrality, objectivity and transparency are conceivably planar. The two-dimensional perspective disguises any general background information in terms of mechanisms (such as algorithms, methodologies used etc.) that leaves the tool as a partial black box. Google itself benefits from this privileged perspective, as it creates the illusion of objectivity. Moreover, users cannot anticipate on the tool’s inner-workings that remain black boxes, which perseveres Google’s authority and the fundamental power relation between Google and its users. Minorities on the other hand, might not benefit from this perspective as shown in figure 6.

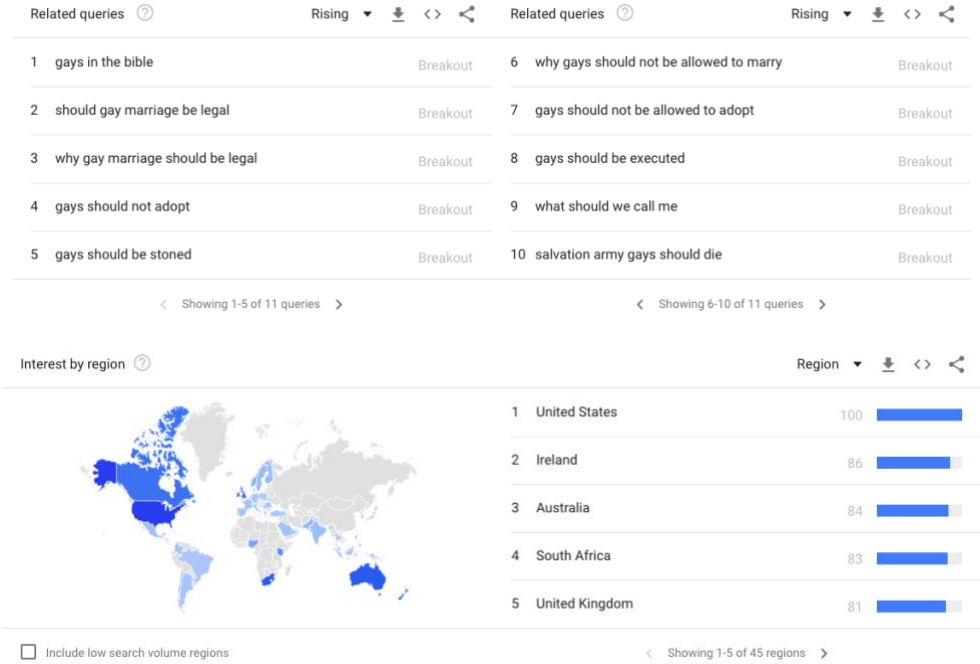


Figure 6. Results *related queries* and *interest by region* on query ‘gays should’ <sup>133134</sup>

The input given by minorities cannot counterbalance the input given by majority groups, such as privileged Americans. The lack of transparency disables minority groups to respond to the tool’s inner-workings in order to highlight their beliefs and advocate their own interests.

Additionally to the epistemic relation between Google Trends and its users, it is arguable that this relation might be augmented with an ontic relation. As the tool presupposes that it serves as an environment in which users can “explore what the world is searching for,” it somewhat implies that the

<sup>133</sup> Keyword entered: “gays should”. Filters: “Worldwide”, “2004-present”, “All categories” and “Web search”. Related queries: “Rising”.

<sup>134</sup> Images acquired from <https://trends.google.com/trends/explore?date=all&q=gays%20should>

world’s inhabitants collectively use Google.<sup>135</sup> As a result, the tool creates a sense of a virtual social reality that reflects the world’s search behavior. However, it is demonstrated that this presumable social reality not only reflects solely 92% of the worldwide searches, but is also demonstrably biased and facilitates as an environment in which echo chambers are constituted (see figure 7).<sup>136</sup>

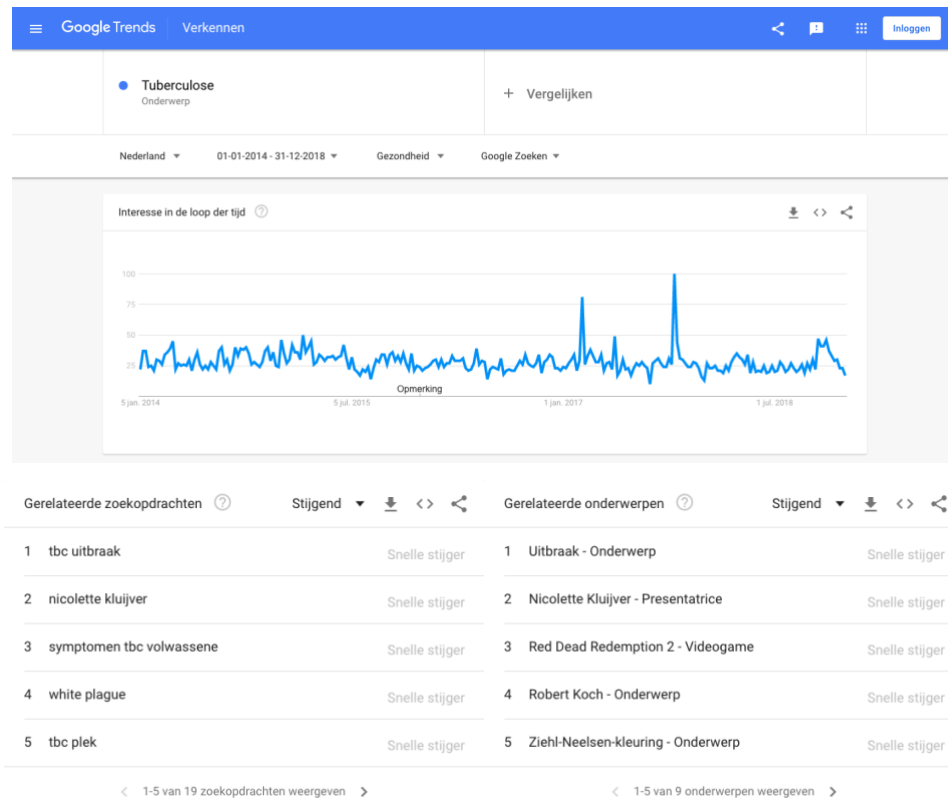


Figure 7. Results on ‘Tuberculose’ (Dutch term for Tuberculosis, translated by author)<sup>137138</sup>

In the case of *interest over time*, one may simply observe the graph and interpret the results as if the number of cases of tuberculosis increased in 2017. While in fact, there were the least tuberculosis patients in The Netherlands in 2017 within the period of 2014 – 2018.<sup>139</sup> The peak in 2017, however, fits the norm in terms of media attention, as the media paid attention to a tuberculosis-outbreak in a Dutch village called Vaassen.<sup>140</sup> Additionally, the apparatuses *related topics* and *related queries* show results on – among others – a Dutch TV host who is somehow linked to the disease. Again, users might interpret these results as if it were the reality in which Nicolette Kluijver is indeed a tuberculosis patient, while the host was actually inaccurately diagnosed and did not suffer from this disease.

<sup>135</sup> Google Trends, “Explore.”

<sup>136</sup> StatCounter, “Search Engine Market Share Worldwide.”

<sup>137</sup> Keyword entered: “Tuberculose”. Filters: “The Netherlands”, “01-01-2014 – 31-12-2018”, “Health” and “Web search”. Related queries: “Rising”.

<sup>138</sup> Images acquired from <https://trends.google.com/trends/explore?cat=45&date=2014-01-01%202018-12-31&geo=NL&q=Tuberculose>

<sup>139</sup> “Tuberculose Kerncijfers Nederland 2018,” Rijksinstituut voor Volksgezondheid en Milieu, accessed June 11, 2019, [https://www.rivm.nl/sites/default/files/2019-03/011225\\_120101\\_Kerncijfers\\_TBC\\_NL\\_V2\\_TG\\_0.pdf](https://www.rivm.nl/sites/default/files/2019-03/011225_120101_Kerncijfers_TBC_NL_V2_TG_0.pdf).

<sup>140</sup> “Honderden inwoners in Vaassen getest op tbc,” NOS, last modified October 24, 2017, <https://nos.nl/artikel/2199472-honderden-inwoners-in-vaassen-getest-op-tbc.html>.

It can be inferred from this that users who do not acknowledge Google Trends' limitations, might interpret the tool's results as if it were the social reality and therefore believe that the tool reflects the whole truth, while these results are not deemed to be objective and neutral. In fact, these results are situated in order to fit certain norms and are presented through the tool's two-dimensional perspective.

## Conclusion

As this research has shown, Google Trends is an exemplary socio-technological system whose epistemic relation with its users is augmented with an ontic relation. As Google Trends presupposes that its users can “explore what the world is searching,” the tool implies that it serves as a virtual environment that reflects an online social reality.<sup>141</sup> The phrase somewhat implies that the world’s inhabitants collectively use Google, and that the collection of search queries entered is displayed in Google Trends. It might be concluded from this that Google Trends creates the lure of a social reality, which is demonstrably considered as non-neutral and situated.

In response to the non-neutrality and situatedness of Google Trends, this research calls for an inquiry that critically depicts on knowledge constitution through other easy accessible socio-technological tools through the framework of Critical Data Studies. As easy accessible socio-technological tools become more ubiquitous, it has now become urgent to apply such frameworks in order to raise awareness on how knowledge constituted through these tools might be biased and limited. However, it is noted that the framework of Critical Data Studies falls short in terms of insights into the users and audiences of tools. Therefore, one can only assume how users might misinterpret and produce limited or biased knowledge through these tools.

The examined apparatuses *terms and comparison*, *filters*, *interest over time*, *interest by subregion*, *related topics* and *related queries* turned out to be deeply intertwined and appeared to collectively create and produce data in the overarching interest of the assemblage – Google Trends – itself. Additionally, by assessing the methodology of Tool Criticism it is shown that the apparatuses individually produce knowledge through a somewhat problematic approach. Users who produce knowledge through *terms and comparison* might receive inappropriate results as they can hardly eliminate the ambiguity of keywords. The results one receives after selecting inappropriate *filters* might be biased, which in turn affects the way in which knowledge is constituted. Additionally, the tool’s visualizations displayed in *interest over time* and *interest by subregion* have shown to create a sense of objectivity and facticity – without being transparent by any means. Subsequently, the elements *related topics* and *related queries* demonstrably contain echo chambers, which means that users produce knowledge based on previous-held beliefs. Furthermore, it is proven that the assemblage itself, as well as its individual apparatuses, are receptive to misinterpretations. Finally, it is shown that the assemblage produces knowledge through a two-dimensional perspective in terms of both interface and inner-workings. Due to the fundamental power relation, Google itself primarily benefits from this perspective, while minority groups consequently suffer from this perspective.

While the results of this research might serve as arguments within the ongoing academic debate on Google Trends objectivity in general, the research cannot contribute to arguments on how the users actually interpret Google Trends’ graphs, maps and other qualitative information, as this research only

---

<sup>141</sup> Google Trends, “Explore.”

assumes how users might interpret the tool's results. Future research must thus identify how users respond to the tool's results through e.g. surveys. Another limitation of this study is its relatively small case studies per apparatus. Future research could incorporate an extensive scope of varying subjects that may substantiate certain claims on Google Trends' non-neutrality and situatedness. The final limitation of this research is that the perspective used is predominantly critical. As a consequence of this perspective, the research might subconsciously focus on the aspects of Google Trends that conform to this perspective.

## References

- Akrich, Madeleine & Bruno Latour. "The De-Description of Technical Objects." In *Shaping Technology / Building Society: Studies in Sociotechnical Change*, edited by Wiebe Bijker & John Law, 205-224. Cambridge, MA: MIT Press, 1992.
- Brey, Philip. "The epistemology and ontology of human-computer interaction." *Minds and Machines* 15, no. 3-4 (2005): 383-398.
- Choi, Hyunyoung, and Hal Varian. "Predicting the present with Google Trends." *Economic Record* 88 (2012): 2-9.
- Dalton, Craig, and Jim Thatcher. "Inflated granularity: The promise of big data and the need for a critical data studies." In *Presentation at the Annual Meeting of the Association of American Geographers, Tampa*. 2014.
- Dalton, Craig, and Jim Thatcher. "What does a critical data studies look like, and why do we care? Seven points for a critical approach to 'big data'." *Society and Space* 29 (2014).
- D'Avanzo, Ernesto, Giovanni Pilato, and Miltiadis Lytras. "Using twitter sentiment and emotions analysis of google trends for decisions making." *Program* 51, no. 3 (2017): 322-350.
- Diakopoulos, Nicholas. "Algorithmic Accountability Reporting: On The Investigation Of Black Boxes." PhD diss., Columbia Journalism School, 2013.
- Flaxman, Seth, Sharad Goel, and Justin M. Rao. "Filter bubbles, echo chambers, and online news consumption." *Public opinion quarterly* 80, no. S1 (2016): 298-320.
- France-Presse, Agence. "Chronology of Taiwan-China Relations Since 1945." *New York Times*, March 18, 2000. <https://archive.nytimes.com/www.nytimes.com/library/world/asia/031800china-taiwan-chrono.html>.
- Golovin, Daniel, Benjamin Solnik, Subhdeep Moitra, Greg Kochanski, John Karro, and D. Sculley. "Google vizier: A service for black-box optimization." In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1487-1495. ACM, 2017.
- Google. "Compare Trends search terms." Trends Help. Accessed March 18, 2019. <https://support.google.com/trends/answer/4359550?hl=en>.
- Google. "Search using autocomplete." Google Search Help. Accessed October 22, 2018. <https://support.google.com/websearch/answer/106230?co=GENIE.Platform%3DAndroid&hl=en>.
- Google Trends. "Explore." Accessed May 2, 2019. <https://trends.google.com/trends/explore>.
- Grömping, Max. "'Echo Chambers' Partisan Facebook Groups during the 2014 Thai Election." *Asia Pacific Media Educator* 24, no. 1 (2014): 39-59.

- Hannak, Aniko, Piotr Sapiezynski, Arash Molavi Kakhki, Balachander Krishnamurthy, David Lazer, Alan Mislove, and Christo Wilson. "Measuring personalization of web search." In *Proceedings of the 22nd international conference on World Wide Web*, pp. 527-538. ACM, 2013.
- Haraway, Donna. "Situated knowledges: The science question in feminism and the privilege of partial perspective." *Feminist studies* 14, no. 3 (1988): 575-599.
- Harding, Sandra. *Objectivity and diversity: Another logic of scientific research*. University of Chicago Press, 2015.
- Heersmink, Richard. "A virtue epistemology of the Internet: Search engines, intellectual virtues and education." *Social Epistemology* 32, no. 1 (2018): 1-12.
- IMDb. "Ching Chong Chang." *Orange Is the New Black*. Accessed June 7, 2019. [https://www.imdb.com/title/tt3807528/trivia?ref\\_=tt\\_trv\\_trv](https://www.imdb.com/title/tt3807528/trivia?ref_=tt_trv_trv).
- Internet Live Stats. "Google Search Statistics." Trends and More. Accessed June 7, 2019. <https://www.internetlivestats.com/google-search-statistics/>.
- Jun, Seung-Pyo, Hyoungh Sun Yoo, and San Choi. "Ten years of research change using Google Trends: From the perspective of big data utilizations and applications." *Technological Forecasting and Social Change* 130 (2018): 69-87.
- Kennedy, Helen, Rosemary Lucy Hill, Giorgia Aiello, and William Allen. "The work that visualisation conventions do." *Information, Communication & Society* 19, no. 6 (2016): 715-735.
- Kitchin, Rob, and Tracey Lauriault. "Towards critical data studies: Charting and unpacking data assemblages and their work." (2014): 3-18.
- Koninkrijk der Nederlanden. "Taiwan." *Nederland wereldwijd*. Accessed June 7, 2019. <https://www.nederlandwereldwijd.nl/landen/taiwan>.
- Koolen, Marijn, Jasmijn van Gorp, and Jacco van Ossenbruggen. "Toward a model for digital tool criticism: Reflection as integrative practice." *Digital Scholarship in the Humanities* 34, no. 2 (2018): 368-385.
- Kristoufek, Ladislav. "Can Google Trends search queries contribute to risk diversification?." *Scientific reports* 3 (2013): 2713.
- Lauriault, Tracey. "Data, Infrastructures and Geographical Imaginations." PhD diss., Carleton University, 2012.
- McLuhan, Marshall. *Understanding Media: The Extensions of Man*. New York: McGraw-Hill, 1966.
- Mellon, Jonathan. "Internet search data and issue salience: The properties of Google Trends as a measure of issue salience." *Journal of Elections, Public Opinion & Parties* 24, no. 1 (2014): 45-72.
- Monsees, David, and Marlo McGriff. "Introducing auto-delete controls for your Location History and activity data." *The Keyword*. Accessed May 9, 2019. <https://blog.google/technology/safety-security/automatically-delete-data/>.

- NOS. "Google verzweeg bug waardoor gebruikersdata ingezien kon worden." Tech. Accessed May 9, 2019. <https://nos.nl/artikel/2253973-google-verzweeg-bug-waardoor-gebruikersdata-ingezien-konden-worden.html>.
- NOS. "Honderden inwoners in Vaassen getest op tbc." Regio. Last modified October 24, 2017. <https://nos.nl/artikel/2199472-honderden-inwoners-in-vaassen-getest-op-tbc.html>.
- NOS. "Klacht tegen Google vanwege 'ontfutselen' locatiegegevens." Tech. Last modified November 11, 2018. <https://nos.nl/artikel/2261042-klacht-tegen-google-vanwege-ontfutselen-locatiegegevens.html>.
- NOS. "Mensen kunnen straks automatisch Google-data laten verwijderen." Tech. Accessed May 9, 2019. <https://nos.nl/artikel/2282875-mensen-kunnen-straks-automatisch-google-data-laten-verwijderen.html>.
- Palos-Sanchez, Pedro R., and Marisol B. Correia. "The collaborative economy based analysis of demand: Study of Airbnb case in Spain and Portugal." *Journal of theoretical and applied electronic commerce research* 13, no. 3 (2018): 85-98.
- Pasquale, Frank. *The black box society*. Harvard University Press, 2015.
- Rijksinstituut voor Volksgezondheid en Milieu. "Tuberculose Kerncijfers Nederland 2018." Accessed June 11, 2019. [https://www.rivm.nl/sites/default/files/201903/011225\\_120101\\_Kerncijfers\\_TBC\\_NL\\_V2\\_TG\\_0.pdf](https://www.rivm.nl/sites/default/files/201903/011225_120101_Kerncijfers_TBC_NL_V2_TG_0.pdf).
- Rogers, Richard. "Foundations of Digital Methods: Query Design." In *The Datafied Society: Studying Culture Through Data*, edited by Mirko Tobias Schäfer and Karin van Es, 75-94. Amsterdam: Amsterdam University Press, 2017.
- StatCounter. "Search Engine Market Share Worldwide." Accessed June 8, 2019. <http://gs.statcounter.com/search-engine-market-share#monthly-201905-201905-bar>.
- Stephens-Davidowitz, Seth, and Hal Varian. "A hands-on guide to Google data." *Mountain View, CA* (2014).
- Sunstein, Cass R. *Echo chambers: Bush v. Gore, impeachment, and beyond*. Princeton, NJ: Princeton University Press, 2001.
- Universiteit van Amsterdam. "dhr. prof. dr. R.A. (Richard) Rogers." Accessed March 24, 2019. <http://www.uva.nl/profiel/r/o/r.a.rogers/r.a.rogers.html?1553443885418>.
- Van Es, Karin, Maranke Wieringa, and Mirko Tobias Schäfer. "Tool Criticism: From Digital Methods to Digital Methodology." In *Proceedings of the 2nd International Conference on Web Studies*, pp. 24-27. ACM, 2018.
- Wolf, Jordan Taylor. "Trending in the Right Direction: Using Google Trends Data as a Measure of Public Opinion During a Presidential Election." PhD diss., Virginia Tech, 2018.
- Zhu, J., Xiaohua Wang, Jie Qin, and Lingfei Wu. "Assessing public opinion trends based on user search queries: validity, reliability, and practicality." In *The Annual Conf. of the World Association for Public Opinion Research*, pp. 1-7. 2012.



# Appendix 1. The looping effect

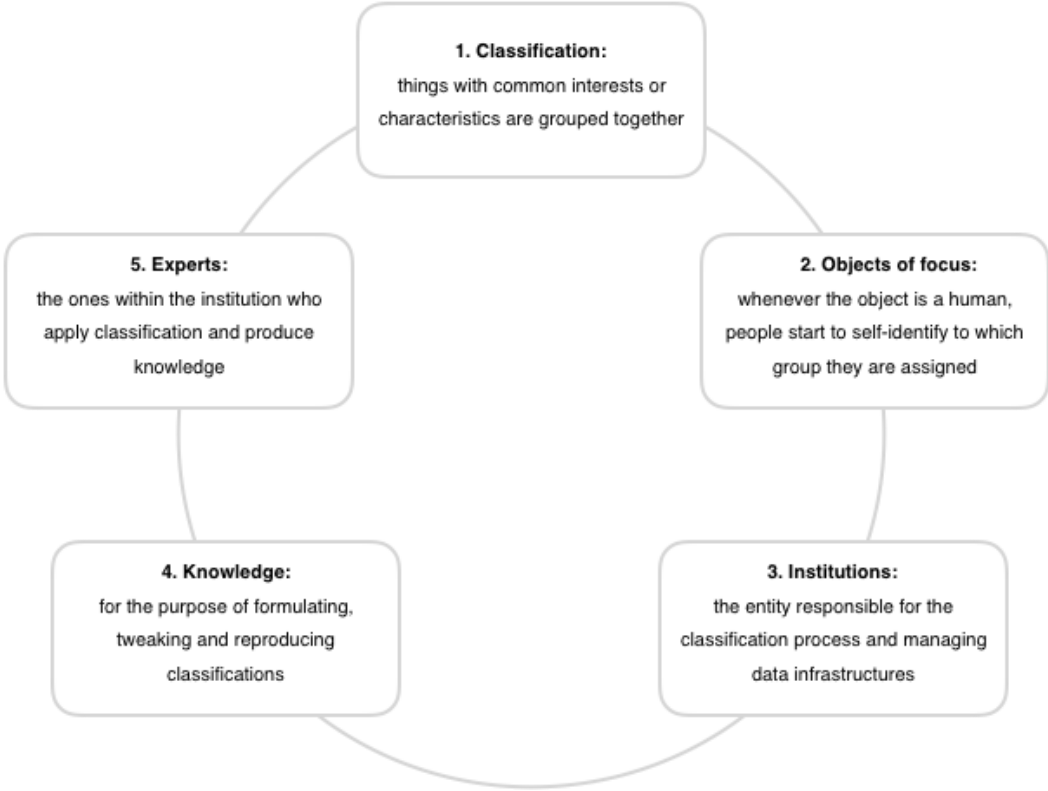


Figure 8. Ian Hacking's looping effect <sup>142</sup>

<sup>142</sup> Kitchin, and Lauriault, "Towards critical data studies," 10-12.

## Appendix 2. Nine intellectual virtues

1. *Curious people* minimize epistemic ramifications because of their genuine motivation to gain a greater understanding of a topic. They wonder and ask why things are the way they are.<sup>143</sup>
2. An *intellectual autonomy* is cognitively capable and thus able to think for oneself. Therefore, this character does not strongly depend on others for acquiring their beliefs.<sup>144</sup>
3. The *intellectual humility* is aware of its personal cognitive limitations and is realistic about knowledge and intellectual domain that are limited.<sup>145</sup>
4. *Attentiveness* is the character of people who are focused when performing cognitive tasks.<sup>146</sup>
5. People with *intellectual carefulness* avoid intellectual errors by means of critical thinking skills.<sup>147</sup>
6. *Intellectual thoroughness* is characterized as the aim for deeper meaning and understanding.<sup>148</sup>
7. *Open-minded people* are epistemologically benefitting from their consideration of alternatives, as they form the most accurate beliefs.<sup>149</sup>
8. *Intellectual courage* is best described as “subjecting ourselves to a potential loss or harm in the context of a distinctively intellectual pursuit like learning or inquiring after the truth.”<sup>150</sup>
9. Finally, the virtue of *intellectual tenacity* is characterized by its determination to pursue its epistemic goals.<sup>151</sup>

---

<sup>143</sup> Heersmink, "A virtue epistemology of the Internet," 3.

<sup>144</sup> Heersmink, "A virtue epistemology of the Internet," 3.

<sup>145</sup> Heersmink, "A virtue epistemology of the Internet," 3.

<sup>146</sup> Heersmink, "A virtue epistemology of the Internet," 4.

<sup>147</sup> Heersmink, "A virtue epistemology of the Internet," 4.

<sup>148</sup> Heersmink, "A virtue epistemology of the Internet," 4.

<sup>149</sup> Heersmink, "A virtue epistemology of the Internet," 4.

<sup>150</sup> Heersmink, "A virtue epistemology of the Internet," 4.

<sup>151</sup> Heersmink, "A virtue epistemology of the Internet," 4.

## Appendix 3. ‘Engines of discoverability’ within Google Trends

Apparatus	Interest over time	Interest by subregion	Related topics	Related queries
<b>Counting</b> Measuring volumes of phenomena	Counts volume over filtered period of time	Counts volume over time per geographical area	Counts volume of related topics over time	Counts volume of related queries over time
<b>Quantifying</b> Turns counts into rates, measurements or classifications	Turns counts into index of 0 – 100	Turns counts into index of 0 – 100	Counts are classified as ‘top’ or ‘rising’ Counts associated with ‘top’ are converted into index of 0 – 100 Counts associated with ‘rising’ are converted into percentages	Counts are classified as ‘top’ or ‘rising’ Counts associated with ‘top’ are converted into index of 0 – 100 Counts associated with ‘rising’ are classified as ‘Breakout’ or converted into a percentage
<b>Creating norms</b> Identify what might be expected	Expectations based on seasonality E.g. hay fever searches in summer > winter	Expectations based on varying province issues E.g. earthquake searches Groningen > Utrecht	Expectations based on entered topic E.g. related topic ‘hay fever’ might be ‘cure’	Expectations based on entered query E.g. related query ‘earthquake Groningen’ might be 112 Groningen
<b>Correlation</b> Define relationships between measurements	By indexing: if value is 50, entered term is half as popular on given timestamp relative to most popular timestamp	By scales: if value is 50, entered term is half as popular in given geographical area relative to the most popular geographical area	By percentages (rising): shows increase in search frequency, and shows ‘Breakout’ when there is a tremendous increase By scales (top): if value is 50, given related topic is half as commonly searched for as related topic with value of 100	By percentage (rising): shows increase in search frequency, and shows ‘Breakout’ when there is a tremendous increase By scales (top): is value is 50, given search query is half as commonly searched for as search query with value of 100
<b>Taking action</b> Employ knowledge to handle issues	Black box	Black box	Black box	Black box
<b>Scientification</b> Provide and follow scientific knowledge	Black box	Black box	Black box	Black box
<b>Normalization</b> Shape the world, so it fits norms	E.g. norms created through media attention on particular timestamp Google can fit norms using proposed related completions	E.g. norms created through media attention on particular geographical area Google can fit norms using proposed related completions	Google has power to fashion the world by means of encouraging certain topics by proposing related topics and thus fit norms	Google has power to fashion the world by means of encouraging certain searches by proposing related topics and thus fit norms
<b>Bureaucratization</b> System to manage and administer people	Black box	Black box	Black box	Black box
<b>Resistance</b> Actions by humans to take back control	Minorities are able to highlight their group/beliefs by collectively entering identical search terms	Minor provinces or countries are able to highlight their beliefs/interests by collectively entering identical search terms.	If minority groups are classified as and connected to certain negative related topics, these humans could forward alternative and more appropriate topics.	Minority groups that are classified as and connected to negative related queries, minority groups could forward alternative and more appropriate topics.

Table 2. ‘Engines of discoverability’ within Google Trends’ data processing apparatuses

## Appendix 4. Example (anti-)programmes and neutral keywords

Figure 9 displays four panels of Google Trends related queries for the keyword 'African American'. Each panel shows a list of queries, their relative ranking (e.g., 'Rising'), and their status (e.g., 'Breakout').

Rank	Query	Status
1	african american museum tickets	Breakout
2	smithsonian african american museum	Breakout
3	african american history museum dc	Breakout
4	short haircuts	Breakout
5	egfr african american	Breakout
6	national african american museum dc	Breakout
7	african american bob hairstyles	Breakout
8	african american museum dc tickets	Breakout
9	hairstyles for african american girls	Breakout
10	museum of african american history dc	Breakout
11	gfr african american	Breakout
12	african american museum hours	Breakout
13	african american history museum tickets	Breakout
14	african american little girl hairstyles	Breakout
15	little girl hairstyles	Breakout
16	natural hairstyles for african american women	Breakout
17	african american men hairstyles	Breakout
18	african american soldiers in the civil war	Breakout
19	african american festival baltimore	Breakout
20	first african american astronaut	Breakout

Figure 9. Results related queries programme and neutral keyword 'African American'<sup>152153</sup>

Figure 10 displays four panels of Google Trends related queries for the keyword 'niggers'. Each panel shows a list of queries, their relative ranking (e.g., 'Rising'), and their status (e.g., 'Breakout' or percentage increase).

Rank	Query	Status
1	niggers in paris	Breakout
2	niggers meme	Breakout
3	ugly niggers	Breakout
4	ship those niggers back	Breakout
5	chris rock niggers	Breakout
6	i hate niggers song	Breakout
7	niggers for sale	Breakout
8	some niggers never die	Breakout
9	lbj niggers	Breakout
10	fuckin niggers	Breakout
11	pack of niggers	Breakout
12	reddit niggers	Breakout
13	how many niggers are in my store	Breakout
14	i smell niggers	Breakout
15	nigger song	Breakout
16	detroit niggers	Breakout
17	ferguson niggers	Breakout
18	fuck niggers	+500%
19	the niggers	+500%
20	black people	+450%

Figure 10. Results related queries anti-programme 'niggers'<sup>154155</sup>

<sup>152</sup> Keyword entered: "African American". Filters: "United States", "2004-present", "All categories" and "Web search".

<sup>153</sup> Images acquired from <https://trends.google.com/trends/explore?date=all&geo=US&q=african%20american>

<sup>154</sup> Keyword entered: "niggers". Filters: "United States", "2004-present", "All categories" and "Web search".

<sup>155</sup> Images acquired from <https://trends.google.com/trends/explore?date=all&geo=US&q=niggers>

## Appendix 5. List countries Google Trends in *filters* <sup>156</sup>

1	Worldwide	63	Denmark	125	Lebanon	187	Samoa
2	Afghanistan	64	Djibouti	126	Lesotho	188	San Marino
3	Åland Islands	65	Dominica	127	Liberia	189	São Tomé & Príncipe
4	Albania	66	Dominican Republic	128	Libya	190	Saudi Arabia
5	Algeria	67	Ecuador	129	Liechtenstein	191	Senegal
6	American Samoa	68	Egypt	130	Lithuania	192	Serbia
7	Andorra	69	El Salvador	131	Luxembourg	193	Seychelles
8	Angola	70	Equatorial Guinea	132	Macao	194	Sierra Leone
9	Anguilla	71	Eritrea	133	Madagascar	195	Singapore
10	Antarctica	72	Estonia	134	Malawi	196	Sint Maarten
11	Antigua & Barbuda	73	Eswatini	135	Malaysia	197	Slovakia
12	Argentina	74	Ethiopia	136	Maldives	198	Slovenia
13	Armenia	75	Falkland Islands (Islas Malvinas)	137	Mali	199	Solomon Islands
14	Aruba	76	Faroe Islands	138	Malta	200	Somalia
15	Australia	77	Fiji	139	Marshall Islands	201	South Africa
16	Austria	78	Finland	140	Martinique	202	South Georgia & South Sandwich Islands
17	Azerbaijan	79	France	141	Mauritania	203	South Korea
18	Bahamas	80	French Guiana	142	Mauritius	204	South Sudan
19	Bahrain	81	French Polynesia	143	Mayotte	205	Spain
20	Bangladesh	82	French Southern Territories	144	Mexico	206	Sri Lanka
21	Barbados	83	Gabon	145	Micronesia	207	St. Barthélemy
22	Belarus	84	Gambia	146	Moldova	208	St. Helena
23	Belgium	85	Georgia	147	Monaco	209	St. Kitts & Nevis
24	Belize	86	Germany	148	Mongolia	210	St. Lucia
25	Benin	87	Ghana	149	Montenegro	211	St. Martin
26	Bermuda	88	Gibraltar	150	Montserrat	212	St. Pierre & Miquelon
27	Bhutan	89	Greece	151	Morocco	213	St. Vincent & Grenadines
28	Bolivia	90	Greenland	152	Mozambique	214	Sudan
29	Bosnia & Herzegovina	91	Grenada	153	Myanmar (Burma)	215	Suriname
30	Botswana	92	Guadeloupe	154	Namibia	216	Svalbard & Jan Mayen
31	Bouvet Island	93	Guam	155	Nauru	217	Sweden
32	Brazil	94	Guatemala	156	Nepal	218	Switzerland
33	British Indian Ocean Territory	95	Guernsey	157	Netherlands	219	Syria
34	British Virgin Islands	96	Guinea	158	New Caledonia	220	Taiwan
35	Brunei	97	Guinea-Bissau	159	New Zealand	221	Tajikistan
36	Bulgaria	98	Guyana	160	Nicaragua	222	Tanzania
37	Burkina Faso	99	Haiti	161	Niger	223	Thailand
38	Burundi	100	Heard & McDonald Islands	162	Nigeria	224	Timor-Leste
39	Cambodia	101	Honduras	163	Niue	225	Togo
40	Cameroon	102	Hong Kong	164	Norfolk Island	226	Tokelau
41	Canada	103	Hungary	165	North Korea	227	Tonga
42	Cape Verde	104	Iceland	166	North Macedonia	228	Trinidad & Tobago
43	Caribbean Netherlands	105	India	167	Northern Mariana Islands	229	Tunisia
44	Cayman Islands	106	Indonesia	168	Norway	230	Turkey
45	Central African Republic	107	Iran	169	Oman	231	Turkmenistan
46	Chad	108	Iraq	170	Pakistan	232	Turks & Caicos Islands
47	Chile	109	Ireland	171	Palau	233	Tuvalu
48	China	110	Isle of Man	172	Palestine	234	U.S. Outlying Islands
49	Christmas Island	111	Israel	173	Panama	235	U.S. Virgin Islands
50	Cocos (Keeling) Islands	112	Italy	174	Papua New Guinea	236	Uganda
51	Colombia	113	Jamaica	175	Paraguay	237	Ukraine
52	Comoros	114	Japan	176	Peru	238	United Arab Emirates
53	Congo - Brazzaville	115	Jersey	177	Philippines	239	United Kingdom
54	Congo - Kinshasa	116	Jordan	178	Pitcairn Islands	240	United States
55	Cook Islands	117	Kazakhstan	179	Poland	241	Uruguay
56	Costa Rica	118	Kenya	180	Portugal	242	Uzbekistan
57	Côte d'Ivoire	119	Kiribati	181	Puerto Rico	243	Vanuatu
58	Croatia	120	Kosovo	182	Qatar	244	Vatican City
59	Cuba	121	Kuwait	183	Réunion	245	Venezuela
60	Curaçao	122	Kyrgyzstan	184	Romania	246	Vietnam
61	Cyprus	123	Laos	185	Russia	247	Wallis & Futuna
62	Czechia	124	Latvia	186	Rwanda	248	Western Sahara
						249	Yemen
						250	Zambia
						251	Zimbabwe

<sup>156</sup> Google Trends, "Explore."

# Appendix 6. Google searches 1999-2012

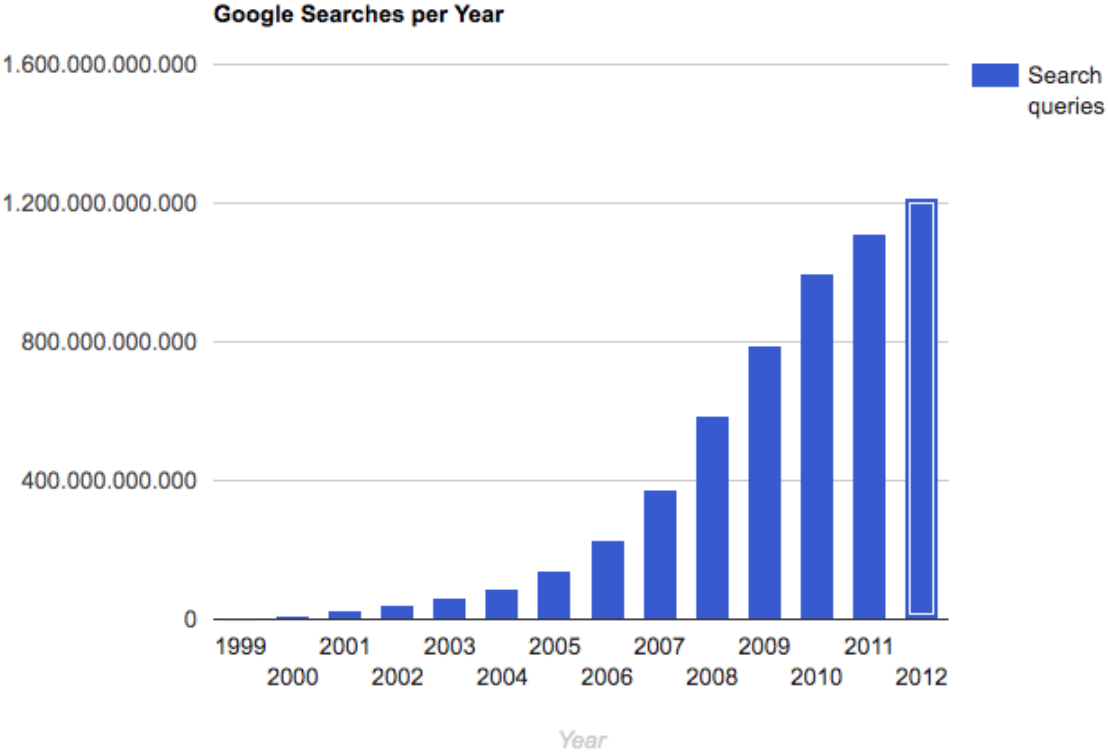


Figure 12. Graph Google searches 1999-2012 <sup>157</sup><sup>158</sup>

<sup>157</sup> Image acquired from <https://www.internetlivestats.com/google-search-statistics/>.

<sup>158</sup> Internet Live Stats, "Google Search Statistics."

## Appendix 7. Google Trends' categories in *filters* <sup>159</sup>

- All categories
- Arts & Entertainment
- Autos & Vehicles
- Beauty & Fitness
- Books & Literature
- Business & Industrial
- Computers & Electronics
- Finance
- Food & Drink
- Games
- Health
- Hobbies & Leisure
- Home & Garden
- Internet & Telecom
- Jobs & Education
- Law & Government
- News
- Online Communities
- People & Society
- Pets & Animals
- Real Estate
- Reference
- Science
- Shopping
- Sports
- Travel

---

<sup>159</sup> Google Trends, "Explore."

# Appendix 8. Results term 'rose' when excluding categories

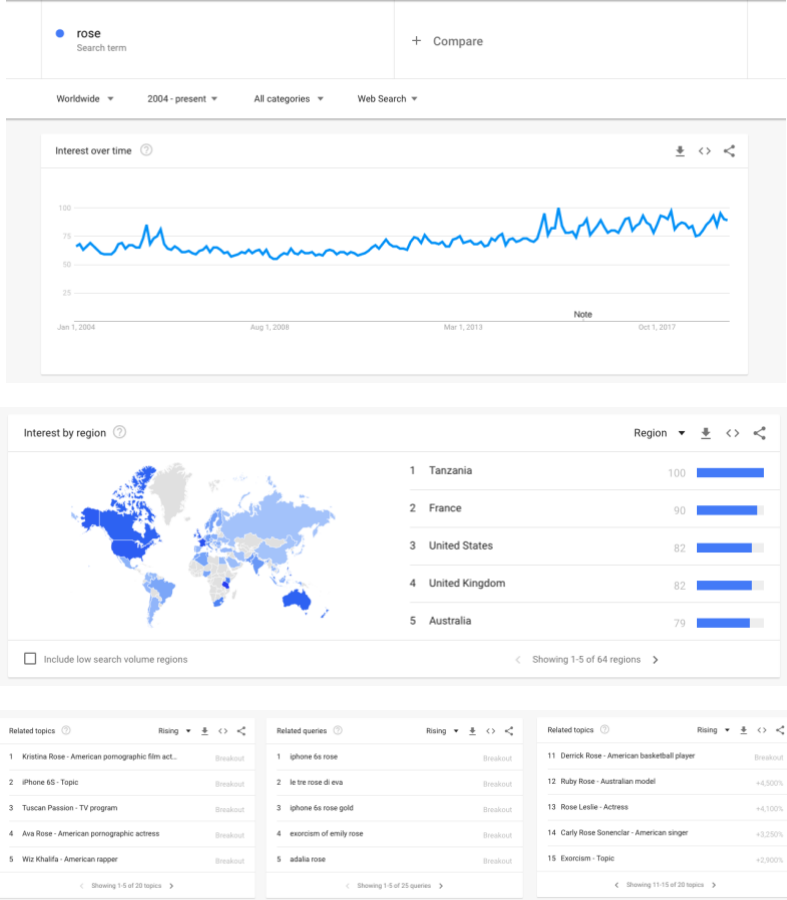


Figure 13. Results term 'rose' without filtering categories <sup>160161</sup>

<sup>160</sup> Keyword entered: "rose". Filters: "Worldwide", "2004-present", "All categories" and "Web search".

<sup>161</sup> Images acquired from <https://trends.google.com/trends/explore?date=all&q=rose>



# Appendix 9. Results term 'rose' when including categories

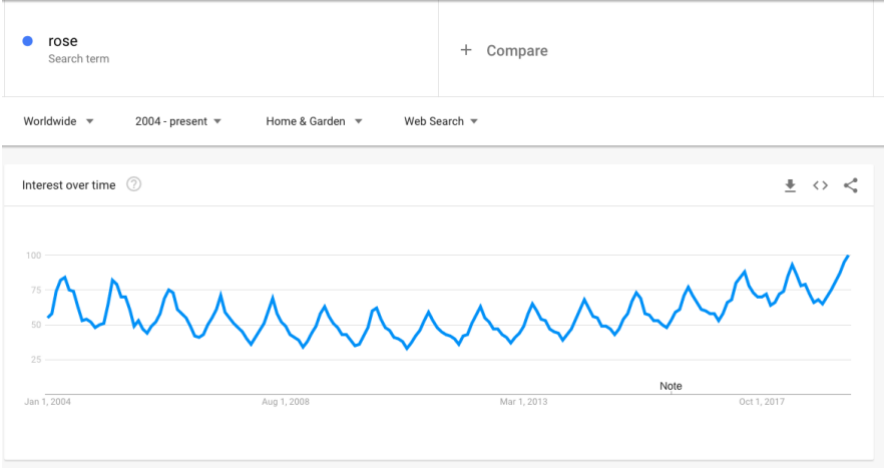


Figure 14. Results *interest over time* combination 'rose' and Home & Garden <sup>162163164</sup>

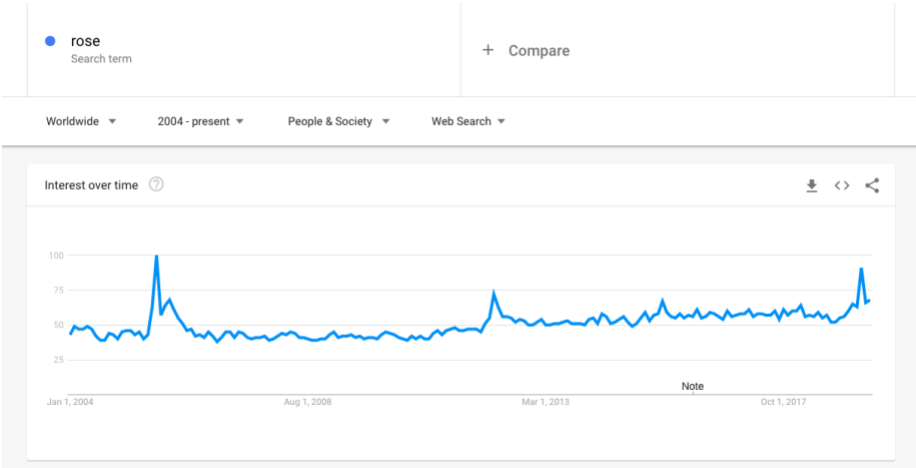


Figure 15. Results *interest over time* combination 'rose' and People & Society <sup>165166167168169</sup>

<sup>162</sup> Keyword entered: "rose". Filters: "Worldwide", "2004-present", "Home & Garden" and "Web search".

<sup>163</sup> Images acquired from <https://trends.google.com/trends/explore?cat=11&date=all&q=rose>

<sup>164</sup> Annual peak in May when roses might be in bloom.

<sup>165</sup> Keyword entered: "rose". Filters: "Worldwide", "2004-present", "People & Society" and "Web search".

<sup>166</sup> Images acquired from <https://trends.google.com/trends/explore?cat=14&date=all&q=rose>

<sup>167</sup> "Ching Chong Chang." Orange Is the New Black, IMDb, accessed June 7, 2019, [https://www.imdb.com/title/tt3807528/trivia?ref\\_=tt\\_trv\\_trv](https://www.imdb.com/title/tt3807528/trivia?ref_=tt_trv_trv).

<sup>168</sup> However, notable that first name is crucial to find appropriate information.

<sup>169</sup> Peak in June 2015, first appearance of actress Ruby Rose in tv-series Orange Is the New Black.

# Appendix 10. Search Engine Market Share Worldwide May 2019

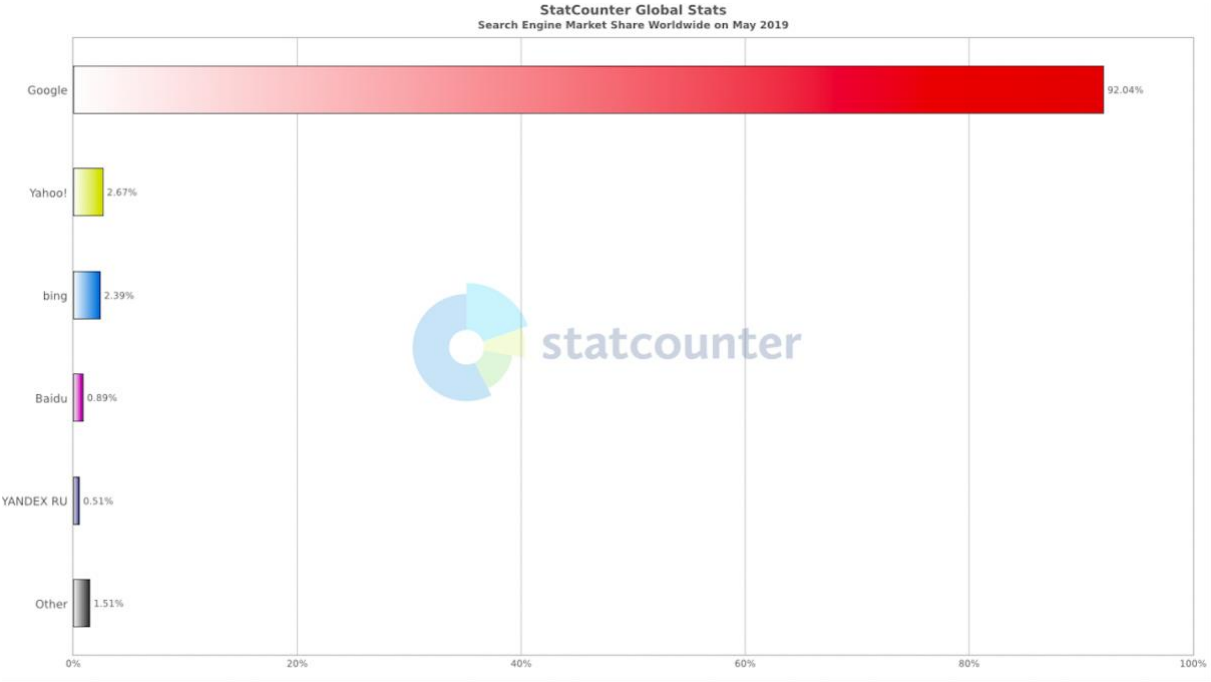


Figure 11. Search Engine Market Share Worldwide on May 2019 <sup>170171</sup>

<sup>170</sup> Image acquired from <http://gs.statcounter.com/search-engine-market-share#monthly-201905-201905-bar>

<sup>171</sup> StatCounter, "Search Engine Market Share Worldwide."

## Appendix 11. Plagiarism awareness declaration form



Universiteit Utrecht

Faculteit Geesteswetenschappen  
Versie september 2014

### VERKLARING KENNISNEMING REGELS M.B.T. PLAGIAAT

#### Fraude en plagiaat

Wetenschappelijke integriteit vormt de basis van het academisch bedrijf. De Universiteit Utrecht vat iedere vorm van wetenschappelijke misleiding daarom op als een zeer ernstig vergrijp. De Universiteit Utrecht verwacht dat elke student de normen en waarden inzake wetenschappelijke integriteit kent en in acht neemt.

De belangrijkste vormen van misleiding die deze integriteit aantasten zijn fraude en plagiaat. Plagiaat is het overnemen van andermans werk zonder behoorlijke verwijzing en is een vorm van fraude. Hieronder volgt nadere uitleg wat er onder fraude en plagiaat wordt verstaan en een aantal concrete voorbeelden daarvan. Let wel: dit is geen uitputtende lijst!

Bij constatering van fraude of plagiaat kan de examencommissie van de opleiding sancties opleggen. De sterkste sanctie die de examencommissie kan opleggen is het indienen van een verzoek aan het College van Bestuur om een student van de opleiding te laten verwijderen.

#### Plagiaat

Plagiaat is het overnemen van stukken, gedachten, redeneringen van anderen en deze laten doorgaan voor eigen werk. Je moet altijd nauwkeurig aangeven aan wie ideeën en inzichten zijn ontleend, en voortdurend bedacht zijn op het verschil tussen citeren, parafaseren en plagiëren. Niet alleen bij het gebruik van gedrukte bronnen, maar zeker ook bij het gebruik van informatie die van het internet wordt gehaald, dien je zorgvuldig te werk te gaan bij het vermelden van de informatiebronnen.

De volgende zaken worden in elk geval als plagiaat aangemerkt:

- het knippen en plakken van tekst van digitale bronnen zoals encyclopedieën of digitale tijdschriften zonder aanhalingstekens en verwijzing;
- het knippen en plakken van teksten van het internet zonder aanhalingstekens en verwijzing;
- het overnemen van gedrukt materiaal zoals boeken, tijdschriften of encyclopedieën zonder aanhalingstekens en verwijzing;
- het opnemen van een vertaling van bovengenoemde teksten zonder aanhalingstekens en verwijzing;
- het parafaseren van bovengenoemde teksten zonder (deugdelijke) verwijzing; parafrazen moeten als zodanig gemarkeerd zijn (door de tekst uitdrukkelijk te verbinden met de oorspronkelijke auteur in tekst of noot), zodat niet de indruk wordt gewekt dat het gaat om eigen gedachtengoed van de student;
- het overnemen van beeld-, geluids- of testmateriaal van anderen zonder verwijzing en zodoende laten doorgaan voor eigen werk;
- het zonder bronvermelding opnieuw inleveren van eerder door de student gemaakt eigen werk en dit laten doorgaan voor in het kader van de cursus vervaardigd oorspronkelijk werk, tenzij dit in de cursus of door de docent uitdrukkelijk is toegestaan;
- het overnemen van werk van andere studenten en dit laten doorgaan voor eigen werk. Indien dit gebeurt met toestemming van de andere student is de laatste medeplichtig aan plagiaat;
- ook wanneer in een gezamenlijk werkstuk door een van de auteurs plagiaat wordt gepleegd, zijn de andere auteurs medeplichtig aan plagiaat, indien zij hadden kunnen of moeten weten dat de ander plagiaat pleegde;
- het indienen van werkstukken die verworven zijn van een commerciële instelling (zoals een internetsite met uittreksels of papers) of die al dan niet tegen betaling door iemand anders zijn geschreven.

De plagiaatregels gelden ook voor concepten van papers of (hoofdstukken van) scripties die voor feedback aan een docent worden toegezonden, voorzover de mogelijkheid voor het insturen van concepten en het krijgen van feedback in de cursushandleiding of scriptieregeling is vermeld.



In de Onderwijs- en Examenregeling (artikel 5.15) is vastgelegd wat de formele gang van zaken is als er een vermoeden van fraude/plagiaat is, en welke sancties er opgelegd kunnen worden.

Onwetendheid is geen excuus. Je bent verantwoordelijk voor je eigen gedrag. De Universiteit Utrecht gaat ervan uit dat je weet wat fraude en plagiaat zijn. Van haar kant zorgt de Universiteit Utrecht ervoor dat je zo vroeg mogelijk in je opleiding de principes van wetenschapsbeoefening bijgebracht krijgt en op de hoogte wordt gebracht van wat de instelling als fraude en plagiaat beschouwt, zodat je weet aan welke normen je je moeten houden.

Hierbij verklaar ik bovenstaande tekst gelezen en begrepen te hebben.
Naam: Charlotte Hanner
Studentnummer: 6562671.
Datum en handtekening: 13-06-2019 Hanner.

Dit formulier lever je bij je begeleider in als je start met je bacheloreindwerkstuk of je master scriptie.

Het niet indienen of ondertekenen van het formulier betekent overigens niet dat er geen sancties kunnen worden genomen als blijkt dat er sprake is van plagiaat in het werkstuk.