# Visualizing missing voices in the history of philosophy: A network analysis of philosophers on Wikipedia

Junyi Tao

## 1 Introduction

The canon of philosophy is traditionally comprised of a select group of esteemed figures and their works, yet beneath this esteemed landscape lie numerous voices that have been historically silenced and systematically excluded from the mainstream narrative. Women, for example, are virtually absent in the canon of philosophy. This assumption that women were incapable of participating in philosophical discussions or contributing to the field is false. In fact, there have been numerous women philosophers who have actively engaged in and significantly contributed to philosophical discourse throughout history (O'Neill 1998). For example, Descartes's renowned creation of the mind-body divide was not solely his own, but was a product of his intellectual collaboration with Princess Elisabeth of Bohemia in their discussion on Meditations on First Philosophy (Pal 2012). Another example is Émilie du Châtelet who was highly renowned during her time and had a significant influence on natural philosophy for her published masterwork The Foundations of Physics and her translations of Newton's Principia (Detlefsen 2018). Those talented women philosophers, though often celebrated in their own time, were later erased from philosophical study through historiographical and institutional efforts. By neglecting their voices, we miss out on valuable conversations and important discussions we could have had, as well as their unique perspectives on various philosophical issues. Furthermore, this lack of inclusivity reinforces the canon-centered narrative in philosophy, exacerbating the problem of underrepresentation and diminishing the diversity necessary for the growth and advancement of the field.

It is widely accepted that there are missing voices in the field of philosophy, but the extent and systematic nature of this problem, as well as its manifestation in the digital age, remains largely unexplored. Traditional methods of close reading, while effective in providing in-depth and nuanced understandings of materials, are labor-intensive and limited in their ability to efficiently analyze large amounts of data and provide a systematic overview. To address these limitations, I employ computational methods, including web crawling, network visualization and analysis, to study the representation and perception of marginalized philosophers on the internet.

This project aims to illustrate and loosen the hold of canonical narrative in philosophy, by visualizing marginalized voices in the network of philosophers and suggesting strategies for increasing the reach and impact of initiatives[1] that seek to highlight these voices. Though mapping and analyzing the connections between philosophers acknowledged on Wikipedia, I will examine the representation of philosophy on the internet through addressing the following questions:

---

[1] Such as Project Vox (https://projectvox.org/), New Narratives in the History of Philosophy (https://www.newnarrativesinphilosophy.net/), and History of Women Philosophers and Scientists (https://historyofwomenphilosophers.org/).

1. Are marginalized philosophers, particularly women philosophers, also depicted as marginalized in this network?

2. If they are marginalized, is this due to a lack of information about them on Wikipedia or the way in which their information is organized?

## 2 Methodology

### 2.1 Data Source

I select Wikipedia as my data source because of its extensive coverage of philosophers and the connections between them, represented by linkages between their pages, and its impacts on both non-academic and academic readers. Students are increasingly relying on search engines and Wikipedia articles to access information online, and scholars, while primarily using specialized searches in academic databases, also frequently use general-purpose search engines such as Google to locate material on specific subjects and may be directed to articles that are frequently cited by Wikipedia and rank high in research results (Meyer and Schroeder 2015).

Additionally, the information on Wikipedia is supposedly less subject to power dynamics and is open to editing by almost anyone, rather than just authorities who may have historically ignored the contributions of women scholars. Given the existence of initiatives featuring marginalized philosophers, as previously mentioned, and the increasing attention paid to them, Wikipedia can reference these sources in a timely manner, and members of these projects are also likely to go edit the corresponding Wikipedia pages of these figures. In this case, are these figures also marginalized on Wikipedia?

### 2.2 Data Preparation

I obtain the names of philosophers from the "list of philosophers"[2] on Wikipedia and then wrote a Python script to access these philosophers' Wikipedia pages. Among these, 66 pages were inaccessible, either because the philosophers only had pages on Wikipedia sites in other languages or because they had multiple name variants. To correct these invalid links, I locate the correct pages of these philosophers and ignore those who are not included in English Wikipedia. Finally, to collect the data, I use DBpedia rather than directly scraping information from Wikipedia pages due to their inconsistent formats, particularly for lesser-known philosophers. DBpedia is a project that extracts structured data from Wikipedia and maintains records of all corresponding philosophers' pages.

### 2.3 Network Construction

With information scraped from these philosophers' pages, I create a network visualization (Figure 1) that maps the linkages between these pages, which represent philosophers' influ-

---

[2]Philosophers are listed by name, and the first page of this list is https://en.wikipedia.org/wiki/List_of_philosophers_(A%E2%80%93C).

Figure 1: The network visualization of philosophers and their influences on each other. Each philosopher is represented as a blue node and their links are in grey.

ences on each other as acknowledged by Wikipedia.[3] This network consists of 1,800 nodes, each representing a philosopher, and features two types of links: "influence" and "influenced by," indicating the direction of influence. For example, Figure 2 illustrates Leibniz's connections with those whom he influenced or who influenced him. Individuals who were influenced by Leibniz and are recognized as such will be listed in the "influenced" section of his Wikipedia page, while those who had an influence on Leibniz will be listed in the "influences" section (see Figure 6). The size of each node in this network is determined by its "importance," calculated using the PageRank algorithm. The more important a node is within the network, the larger it appears. The layout of the nodes reflects their connectivity as nodes that are more connected tend to cluster together.

## 2.4 Identify marginalization from the network

Marginalized philosophers, represented in the network visualization as having few connections to other nodes, can be identified based on both their locations in the network and the sizes of their nodes. The layout of the network is determined by the force-directed algorithm, which simulates the forces acting on each node and determines their positions based on the strength of their connections. Marginalized nodes are typically located further from the center of the network due to their weaker connections to more central nodes. Extreme marginalization can result in complete isolation, represented in the visualization as small, isolated dots around the main network.

The size of a node represents its importance, as measured by the PageRank algorithm, which considers the number of links from other nodes to a given node as an indication of its influence. Philosophers with a high PageRank score, indicated by larger node sizes, are considered more influential within the network. The rank of a philosopher is also influenced

---

[3]The interactive visualization is live on https://www.junyitao.com/research/storyplus2022/InfluNet.
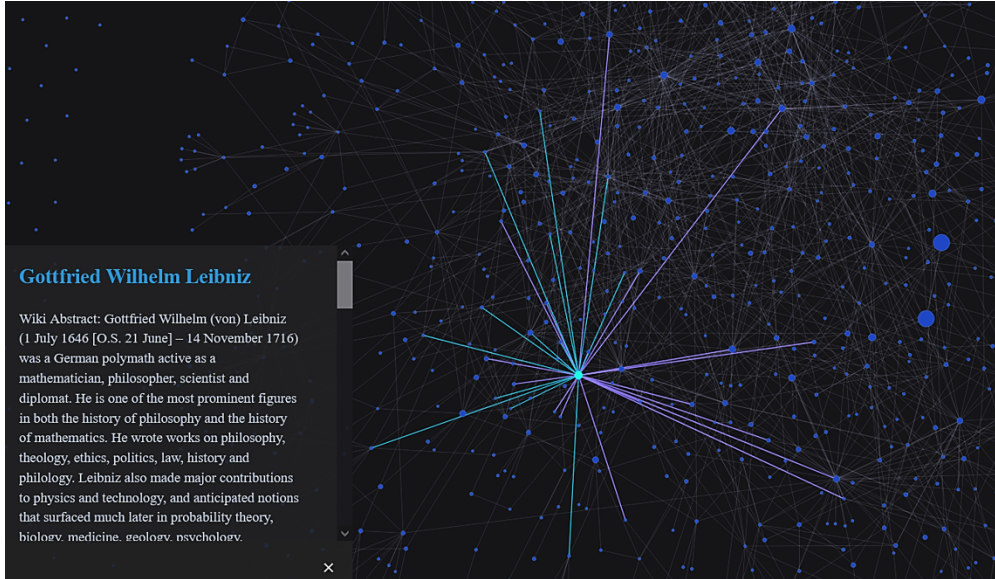
Figure 2: A screenshot of the network with Leibniz selected. Links in light blue represent Leibniz was "influenced by" other philosophers, and links in purple represent Leibniz's "influence" on others.

by the ranks of those they are connected to, so those with many connections to high-ranking philosophers will themselves have a higher rank. Therefore, marginalized philosophers can be identified by looking for those with lower PageRank scores, indicated by smaller node sizes.

## 3 Results and Discussion

### 3.1 Identify marginalized women philosophers in the network

In the network (Figure 1), I identify women philosopher based on the pronouns (she/her) in the descriptions in their Wikipedia pages, and then highlight them in the network visualization. As shown in Figure 3, women philosophers are often only linked to a few other nodes or are completely isolated within the network.

Why are women philosophers so scarce and marginalized in the network? Is this due to a lack of information about them on Wikipedia, or is there another reason? My further investigation of this phenomenon suggests that the issue goes beyond a lack of knowledge about these philosophers. In this section, I analyze two key elements in the network: nodes, representing "who are considered philosophers", and links, representing "whose contribution to or engagement in philosophy is acknowledged". I find great discrepancies in both nodes and links. The results indicate that many women philosophers are not considered "philosophers" in the network, and many of their connections with their contemporaries are missing.
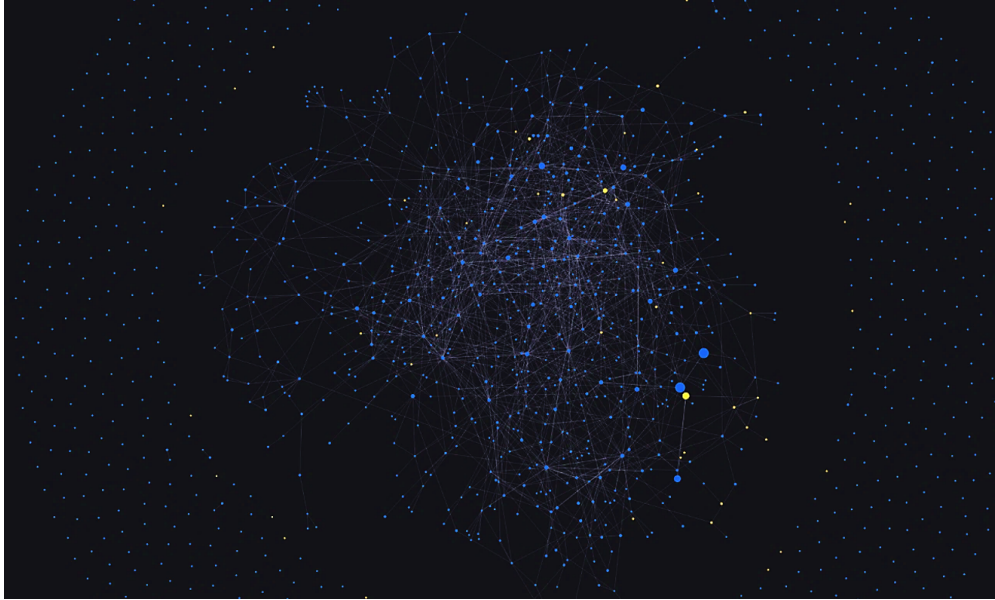
Figure 3: Network of philosopher, with women philosophers highlighted in yellow and other nodes in blue.

### 3.1.1 Missing nodes: Discrepancies in "Philosopher" and "Women Philosopher" List

This network in question contains just 49 female philosophers, which is significantly fewer than the number of women philosophers with entries on Wikipedia. Upon investigation, I discover that there are two distinct lists on a Wikipedia page[4] that compiles collections of philosophers' entries: one for "philosophers" and another for "women philosophers" (see Figure 4). As of May 28, 2022, there are 261 women philosophers listed on the "list of women philosophers" page, but only 49 of them are included in the general list of "Philosophers." This means that 81% of the women philosophers listed on the "list of women philosophers" page are excluded from the narrative of "who is counted as a philosopher." To address this issue, I add female philosophers who are listed exclusively under the "Women Philosophers" list to the network (see Figure 3), resulting in a significant increase in the number of highlighted nodes representing women philosophers.

### 3.1.2 Missing links: Discrepancies in influence links

Even after incorporating previously omitted female philosophers into the network, the majority of them remain disconnected and isolated, with the network remaining fragmented and only the central nodes interconnected. This lack of connectivity is incongruous with the collaborative and discursive nature of philosophy.

Upon examining the links, I discover that there are 1776 "influence" links and 2580

---

[4]There is a page on Wikipedia that provides multiple lists of philosophers (https://en.wikipedia.org/wiki/Lists_of_philosophers), where has a general list of philosophers "lists of philosophers by name" and a list of women philosophers (https://en.wikipedia.org/wiki/List_of_women_philosophers). Accessed May 28, 2022.
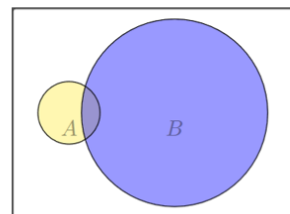
Figure 4: List of philosophers (by name), list of women philosophers, and counts of distinct names in the two lists. Accessed May 28, 2022.

"influenced by" links, but these are meant to be pairs with equal numbers. This suggests that numerous connections between philosophers may have been overlooked due to this discrepancy in links, particularly between mainstream and marginalized philosophers. For example, as Emilie du Châtelet is recognized as a key "contributor to the Leibniz-Newton debate" (Hagengruber 2011), her Wikipedia page is linked to those of Newton and Leibniz, indicating that she was "influenced by" their thoughts. However, there is no link from Leibniz's page to du Châtelet's, indicating that he "influenced" her (see Figure 5). In other words, du Châtelet is not mentioned on Leibniz's Wikipedia page (see Figure 6).

It is understandable that the influences of philosophers may not be consistently updated on Wikipedia, especially for highly influential figures like Leibniz. There might be more people to be recognized as being influenced by him, while not necessarily being recognized on his page. However, my focus is not on ensuring that the pages of these prestigious philosophers are as comprehensive as possible, but rather on those philosophers who are isolated in the network due to missing links. Readers of Wikipedia, which is widely regarded as the most commonly used digital encyclopedia, often expect a certain level of structure and consistency, and may not necessarily check for links that exist in one direction. If marginalized philosophers' pages are not connected to others, even though they should be and could be, their existence becomes invisible to Wikipedia readers, despite the fact that they have their own pages within the encyclopedia.

There are several ways in which the discrepancies in linkages between philosophers' pages and the resulting fragmentation reinforce preexisting biases against female philosophers and other marginalized figures. First, for readers who are non-experts in philosophy and start from the pages of canonical philosophers, they may never be led to the pages of marginalized figures due to missing links. Second, even if readers do discover the pages of these female philosophers, they may find that these philosophers are not linked to any other philosophers, implying that women, despite producing philosophical works, have not truly contributed to the development of philosophy. However, this is not true, as women philosophers, particularly in the early modern period, had extensive connections and intellectual collaborations with
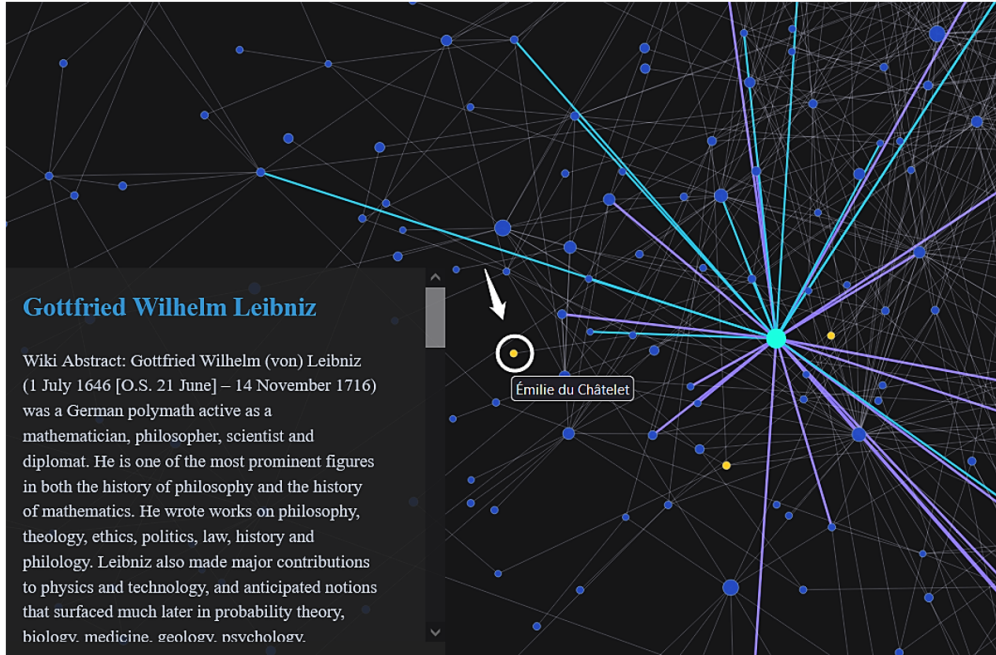
Figure 5: Influence links between Leibniz (the node in light blue) and other philosophers, with no links to du Châtelet (the node pointed by the white arrow).

their contemporaries through correspondence, translation of philosophical works, and the cultivation of friendships (Pal 2012).

These issues can be addressed by re-organizing the linkages between philosophers' Wikipedia pages to ensure that for every link representing "A influenced B," there is a corresponding link indicating "B was influenced by A." For instance, Leibniz and du Châtelet will be linked to each other (Figure 7). Additionally, adding missing links to prestigious philosophers' pages may increase the likelihood of readers accessing the pages of marginalized philosophers, and highlight the connections and contributions of female philosophers whose pages previously showed no connections or influences, reintegrating them into the network.



Figure 6: Screenshots of du Châtelet and Leibniz's Wikipedia pages. People who have influenced them are listed in the "influences" section, and those they influenced are mentioned in the "influenced" section. Accessed May 28, 2022.
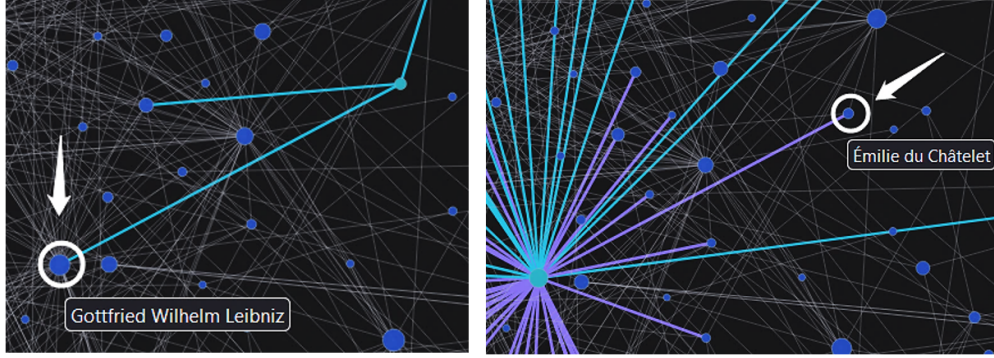
Figure 7: Leibniz and du Châtelet are connected by two directional links, one representing Leibniz "influenced" du Châtelet (left, in light blue) and the other one representing du Châtelet was "influenced by" Leibniz (right, in purple).

## 3.2 Update the network by correcting discrepancies

As shown in Table 1, I recover a total of 3,714 missing links, almost doubling the number of links between philosophers. This correction was particularly important for marginalized women philosophers, who had fewer links compared to main figures. By reconnecting 407 previously isolated philosophers, including 26 women philosophers, to the network, I make these previously marginalized philosophers much more "weaved into" the network. They become more connected with others and move closer to the center of the network (see Figure 8).

Moreover, the landscape of philosophy presented by the network changes significantly once these missing links are recovered, reflecting a more accurate depiction of how philosophy is developed through collaborations between thinkers. This updated network demonstrates that marginalized philosophers have participated in philosophical discourse throughout history, and by acknowledging their role in philosophy, it can increase public interest in and further investigation of their contributions. This creates a positive feedback loop, as the more their works are known to the public, the more likely their Wikipedia pages will be improved and connected to other philosophers, resulting in an even more equitable representation of philosophy.

|  | All philosophers | | | Women philosophers | | |
|---|---|---|---|---|---|---|
|  | Before | Update | After | Before | Update | After |
| Number of nodes | 2012 | - | - | 261 | - | - |
| Number of "influence" links | 1776 | + 2259 | 4035 | 56 | + 40 | 96 |
| Number of "influenced by" links | 2580 | + 1455 | 4035 | 232 | + 75 | 307 |
| Philosophers with no "influence" links | 1599 | - 310 | 1289 | 238 | - 14 | 224 |
| Philosophers with no "influenced by" links | 1398 | - 298 | 1100 | 192 | - 20 | 172 |
| Isolated philosophers | 1365 | - 407 | 958 | 189 | - 26 | 163 |

Table 1: Statistics of philosopher nodes before and after adding missing links.

Figure 8: Network of philosophers with missing links recovered.

## 3.3 Discussion

It is not Wikipedia itself that generates these biases, but rather, it reflects the biases that are prevalent in the representation of philosophy on the internet and how this impacts public understanding. This narrow and biased representation of philosophy creates a negative feedback loop, leading to fewer people from underrepresented groups, such as women and people of color, pursuing careers or further studies in philosophy.

This study demonstrates the importance of distant reading, a method that allows for large-scale analysis of texts and identification of inconsistent links. It also illustrates how traditional close reading practices, if solely focused on canonical figures, can perpetuate the exclusion of historically marginalized voices from philosophical discourse. While contemporary researchers no longer dismiss the philosophical works of women and people of color, they may still lack access to these works. For instance, when students are asked to write about philosophical concepts such as "passion" or "soul," they may only consult the writings of well-known figures, unaware of Princess Elisabeth's extensive engagement with these themes in her correspondences with Descartes.

This research also proposes strategies for increasing the visibility and accessibility of marginalized philosophers on the internet more efficiently. Currently, Wikipedia pages of marginalized philosophers are largely isolated and may only be accessed by those already aware of these philosophers. This limits the reach and impact of initiatives that focus on featuring these marginalized voices and reinforces underrepresentation and exclusion in the current presentation of philosophy online. To address this, initiatives featuring marginalized philosophers can organize researchers to edit links on Wikipedia and other knowledge-sharing platforms, adding connections from well-known philosophers to marginalized figures

to improve access to their pages and increase awareness of the initiatives. Wikipedia and similar platforms can also create a new section providing suggestive links to entries related to marginalized philosophers, increasing their representation and reducing underrepresentation and exclusion.

# 4   Conclusion and Limitation

This project conducts a systematic investigation into the presentation of philosophy on the internet, shedding light on the pervasive problem of marginalization in the field. By analyzing the fragmentation in the network of philosophers on Wikipedia, it reveals how the canon-centered narrative persists and can be reinforced in the digital world, through the sharing and spreading of knowledge online. It also highlights the significance of initiatives that work to feature underrepresented figures and suggests ways for them to more efficiently challenge the exclusive narrative. In summary, it clearly visualizes and suggests ways to loosen the hold of canonical figures in philosophy, which can be applied to other domains of knowledge.

Methodologically, this study illustrates how computational methods can be employed to enhance humanities scholarships in a more reliable and interpretable way, by integrating both close and distant reading methods. As Van Den Berg et al. (2018) noted, it is difficult to determine the accuracy of visualization results due to the potential for errors in coding and corruption of data. For example, the initial network visualization, which omits 81% of female philosophers and has numerous missing links, only further reinforced biased perceptions of philosophy (Figure 3). To address this interpretive issue, this research leverages distant reading to systematically identify problems and close reading to interpret the results and delve deeper into specific cases. In this way, it takes network visualizations not as the final result, but as a starting point to identify and critically examine biases embedded in the sources being used. This also suggests a computation-powered approach to better understand "the sociology of online knowledge" (Meyer and Schroeder 2009), that is, how the dissemination and access to online knowledge is shaped by the socio-technical system.

There are several limitations to be noted. One such limitation is the lack of a clear, standardized definition of "influence" among Wikipedia editors. It is possible that philosophers may exert influence on one another in a variety of ways, and such influence may be subject to debate. Additionally, the mechanisms by which Wikipedia pages are edited and maintained may impact the inclusion of certain philosophers, as suggested by research indicating that pages about women are more likely to be deleted (Tripodi 2021). This raises the possibility that marginalized philosophers such as Émilie du Châtelet may have been mentioned on Leibniz's page but later suggested for deletion, thereby escaping the purview of this study. It would be valuable for future research to delve into the processes and decision-making that go into editing and deleting pages on Wikipedia, possibly through an examination of page editing history.

# References

Detlefsen, Karen (2018). "Émilie du Châtelet". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2018. Metaphysics Research Lab, Stanford University.

Hagengruber, Ruth (2011). *Emilie du Châtelet between Leibniz and Newton*. Vol. 205. ISBN: 978-94-007-2074-9. DOI: 10.1007/978-94-007-2093-0.

Meyer, Eric T and Ralph Schroeder (2009). "Untangling the web of e-Research: Towards a sociology of online knowledge". In: *Journal of informetrics* 3.3, pp. 246–260.

— (2015). *Knowledge machines: digital transformations of the sciences and humanities*. Mit Press.

O'Neill, Eileen (1998). *Disappearing ink: Early modern women philosophers and their fate in history. Philosophy in a feminist voice: Critiques and reconstructions*.

Pal, Carol (2012). *Republic of women: Rethinking the Republic of Letters in the seventeenth century*. Vol. 99. Cambridge University Press.

Tripodi, Francesca (2021). "Ms. Categorized: Gender, notability, and inequality on Wikipedia". In: *New Media & Society*. DOI: 10.1177/14614448211023772.

Van Den Berg, Hein et al. (2018). "A philosophical perspective on visualization for digital humanities". In: *2018 Workshop on Visualization for Digital Humanities*.