Fact Factories:
Wikipedia and the power to represent

Heather Ford
Kellogg College
Word count: 81,209

August 2015

Thesis submitted in partial fulfilment of the requirements for the degree of DPhil in Information, Communication, and the Social Sciences in the Oxford Internet Institute at the University of Oxford.

Acknowledgements

The previous page lists only one name, but so many people were part of the journey that brought this thesis to fruition.

Abstract

Wikipedia is no longer just another source of knowledge about the world. It is fast becoming a central source, used by other powerful knowledge brokers like Google and Bing to offer authoritative answers to search queries about people, places and things and as information infrastructure for a growing number of Web applications and services. Researchers have found that Wikipedia offers a skewed representation of the world that favours some groups at the expense of others so that representations on the platform have repercussions for the subjects of those representations beyond Wikipedia's domain. It becomes critical in this context to understand how exactly Wikipedia's representations come about, what practices give rise to them and what socio-technical arrangements lead to their expression.

This ethnographic study of Wikipedia explores the values, principles and practices that guide what knowledge Wikipedia represents. It follows the foundational principles of Wikipedia in its identity both as an encyclopaedia and a product of the free and open source software and internet freedom rhetoric of the early 2000s. Two case studies are analysed against the backdrop of this ideology, illustrating how different sets of actors battle to extend or reject the boundaries of Wikipedia, and in doing so, affect who are defined as the experts, subjects and revolutionaries of the knowledge that is taken up.

The findings of this thesis indicate that Wikipedia's process of decision-making is neither hierarchical nor is it egalitarian; rather, the power to represent on Wikipedia is rhizoid: it happens at the edges rather than in the centre of the network. Instead of *everyone* having the same power to represent their views on Wikipedia, those who understand how to perform and speak according to Wikipedia's complex technical, symbolic and policy vocabulary tend to prevail over those who possess disciplinary knowledge about the subject being represented. Wikipedians are no *amateurs* as many would have us believe; nor are they passive collectors of knowledge held in sources; Wikipedians are, instead, active co-creators of knowledge in the form of facts that they support using specially chosen sources.

The authority of Wikipedia and Wikipedians is garnered through the performative acts of citation, through the ability of individual editors to construct the traces that represent citation, and through the stabilization and destabilization of facts according to the ideological viewpoints of its editors. In venerating and selecting certain sources among others, Wikipedians also serve to reaffirm traditional centres of authority, while at the same time amplifying new centres of knowledge and denying the authority of knowledge that is not codified in practice. As a result, Wikipedia is becoming the site of new centres of expertise and authoritative knowledge creation, and is signalling a move towards the professionalization of the expertise required to produce factual data in the context of digital networks.

Contents:

Figures:

Tables:

Chapter 1: Introduction

Source: https://twitter.com/Pawelotti/status/609718088242606082

My interest in Wikipedia began when I heard of a case in which Wikipedia editors had repeatedly deleted an article about the Kenyan superhero character, Makmende from English Wikipedia. Makmende is the Sheng (Swahili slang) word for 'hero' that originated from the character Dirty Harry played by Clint Eastwood in the 1983 film 'Sudden Impact'. The term, 'Makmende' (an amalgam of phrase, 'Make my day' uttered by Eastwood's character in the film) was popular in the 1990s in Kenya but enjoyed a resurgence when Nairobi-based band, 'Just a Band' featured Makmende in their YouTube music video for the song, 'Ha He' in March 2010.[1] The music video went viral in Kenya and inspired a series of remixes on Facebook and Twitter. Ethan Zuckerman (2010) explained how Kenyans had tried to start an article dedicated to Makmende on Wikipedia and that they had their contributions repeatedly deleted and I wrote a follow-up to the deletion story in an article about Kenyans' motivations to participate in Wikipedia (Ford, 2011).

My article prompted debate on Wikipedia forums. Through this, and through further research into the practice of article deletions, I learned that my initial reaction to the

---

[1] See https://www.youtube.com/watch?v=_mG1vIeETHc.

case was misguided. I originally believed that ideological interests were the primary force behind the actions taken by editors to remove the article from Wikipedia. Wikipedia editors, being predominantly white, Western men, were simply unwilling to recognise the importance of cultural phenomena taking place far away from them.

Learning more about the process of article deletion, however, I began to recognise two key factors that I had not previously understood. The first was that the actions taken by editors to delete the article were the actions of a very small number of individuals acting on their own interpretation of Wikipedia policy, rather than the entire body of editors interpreting policy in the same way. What I originally recognised as 'Wikipedia deleting' Makmende, I came to understand as a few, widely distributed editors interpreting policy in particular ways that would allow them to legitimise their deletion of the Makmende article, but by extension many other editorial acts as well.

The second was that Wikipedia's processes are not solely human activities carried out on a neutral platform, but are strictly mediated by software code and that these interactions between humans and code follow a predetermined logic that strongly shapes how phenomena within articles are accepted. When the Makmende article was first created it was evaluated by editors in a semi-automated process that employs particular heuristics and filters to determine whether the content met Wikipedia standards. When content moves through Wikipedia's socio-technical system, it is transposed by different work groups, each focusing on particular aspects and using different technologically mediated lenses with which to evaluate it. The representation of Makmende on Wikipedia, in other words, was decontextualized and broken up into myriad different pieces through the semi-automated process that prevented editors

recognising what this knowledge represented for those who were attempting to have it seen by the world.

Furthermore, intricate details about the process were exposed by the debate on the Wikimedia-l mailing list which suggested numerous policy reasons why the article could have been deleted, and equally numerous reasons why it could have been accepted. There was already a Wikipedia article about 'Just a Band'. According to some editors, the best place for facts about Makmende was in that article. On the other hand, it was argued that the Makmende character and the music video that was inspired by him were significant to Kenya's cultural history since they represented Kenya's first Internet meme. In a first attempt at creating an article about Makmende, the author had written a sentence that looked to be vandalism since it took the form of the Chuck Norris-style jokes that emerged from the Makmende meme. The second attempt to create the article constituted a copyright violation because its content, although factual, was taken from a music website and was unattributed. According to editors contributing to the discussion on Wikimedia-l, however, Wikipedians are encouraged to improve articles that are weak, rather than to dismiss them entirely through deletion.

What was clear throughout this passionate debate was that it was really important to Kenyans to have Makmende represented on English Wikipedia. When I interviewed Kenyan Wikipedians in 2011, they saw the deletions as just another series of actions taken by Westerners to dismiss a fact of Kenyan culture as unimportant and un-notable. For the Kenyans whom I spoke with, Wikipedia had originally promised an opportunity for them to have facts about their world displayed equally with others on a global platform. What was originally recognised as an opportunity for recognition, to be seen as a creative participant in a world, became a disappointment through the example of Makmende and other examples of deletionism that I have written about.

More than anything, this event demonstrated how English Wikipedia is being seen as an authoritative platform for facts about the world, that English Wikipedia has significant power to represent knowledge and that Wikipedia editors have become a significant power brokers for the dissemination of authoritative knowledge. It therefore becomes important to understand how Wikipedia filters knowledge about the world. The answer to this question requires more than simplistic theories about individuals with idiosyncratic and unpredictable interests. Instead, Wikipedia, is highly complex, distributed and mediated and the power to represent one's knowledge within Wikipedia's socio-technical system is a result of the mastery of both its technical and social features.

This thesis engages with the question of how an ostensibly egalitarian platform for creating a reliable reference source has reconfigured notions of expertise, authority and the power to represent knowledge in the age of the network. Specifically, it is an attempt to understand the impact of Wikipedia's representational system on who are considered the experts, the research subjects and the revolutionaries of the facts that it represents. Before suggesting how such questions might be answered, it is important to understand exactly how and why Wikipedia has become an important venue for the representation of facts about the world. In order to understand this, we must first look at how Wikipedia has become a critical feature of the infrastructure of the Internet as a whole, particularly in its relationship to search engines such as Google.

1.1   Wikipedia becomes authoritative

On the 16th of May 2012, Google announced a new project called the 'Knowledge Graph' that signalled the move towards a new trend in search. In a blog post entitled

'Introduction to the Knowledge Graph: things not strings', Senior Vice President of Google Engineering, Amit Singhal wrote that Google would be using 'public sources such as Freebase, Wikipedia and the CIA World Factbook' to enable more efficient resolution of queries by users within Google's domain rather than them having to navigate away (Singhal, 2012). Instead of offering the user a long list of possible answers to their queries, Google would also present facts about a user's query in a summarized infobox. A search for 'London', for example, would result in a prominent infobox on the right-hand side of the page listing facts about London extracted from Wikipedia, among other sites as can be seen in Figure 1.1.
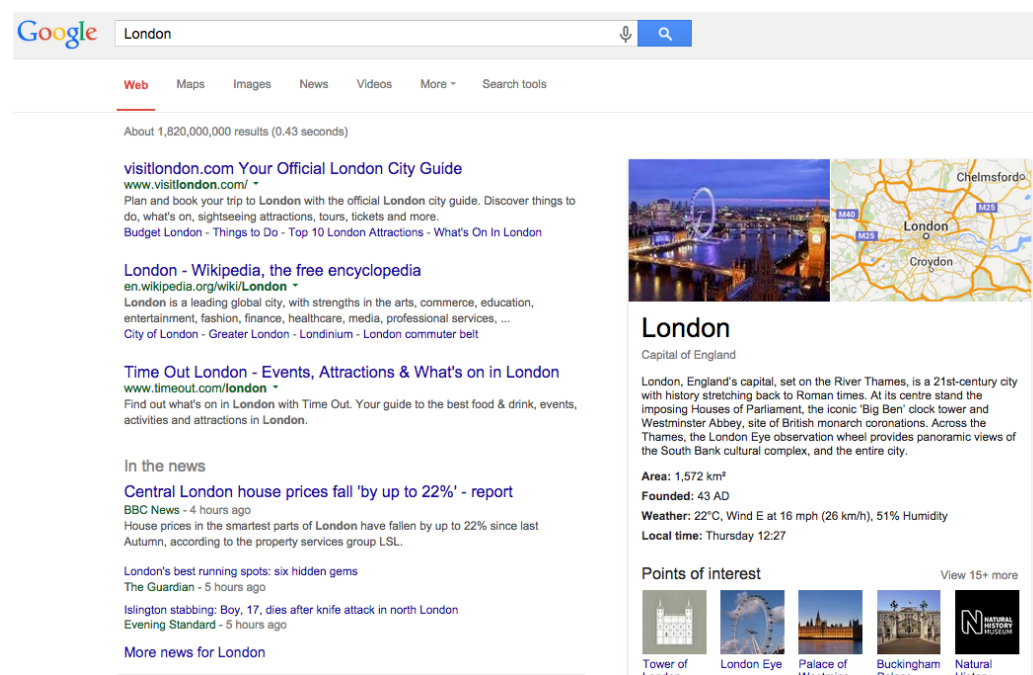
Figure 1.1 Screenshot of results of user query for 'London' on Google

Source: Google.com, 11 June 2015

This move was significant for two key reasons. Firstly, it established Google as a source of facts rather than an indexer of unverified information, and secondly, it validated Wikipedia as an authoritative source of those facts. Instead of Google presenting a list of alternative sources that may provide the answer to a user's query, Google was making a

15

statement that it could also know *the answer* to that query. With Singhal's comment that Wikipedia was a 'public source' of information, Google implied that Wikipedia was representative of a collective consensus that made it authoritative as a source of facts (rather than mere claims) about the world.

Google had always prioritized Wikipedia in its search results. According to a 2012 study by search engine optimisation company Intelligent Positioning, Wikipedia pages appear on the first page of Google for 99% of searches and Wikipedia is the first result on Google for 56% of searches (Silverwood-Cope, 2012). Google has been an important ally to Wikipedia, sending 61% of its traffic to Wikipedia between 2003 and 2008 – a period during which Wikipedia traffic grew by 8,000% (Johnson, 2008).[2] The Knowledge Graph was different, however, because it changed the fundamental communicative relationship between the user, Google and Wikipedia. Instead of a user asking the search engine for all the possible venues in which they could find an answer to their query, the user was being presented with *the* answer and there was an assumption that Wikipedia was the authoritative source of such answers.

Since 2012, other search engines such as Bing, AOL and Lycos have replicated this service by using Wikipedia and other knowledge bases to present facts about a user's query. Facebook similarly presents information from Wikipedia about people, places and things using its Graph search tool (see figure 1.2) and multiple software applications are starting to extract facts from Wikipedia in order to enrich their services. Such extraction activities have been enhanced by a project of the Wikimedia Foundation called Wikidata started in 2012. Funded in part by Google, Wikidata is a database that

---

[2] It isn't only Google that prioritizes Wikipedia in search results. Microsoft's search engine, Bing, was found to favour Wikipedia in search results even more than Google (Goodwin, 2012).

16

extracts information from Wikipedia and other Wikimedia Foundation projects, as well as other sources of open data, and makes it available in both a human and machine-readable format. Wikidata's format and open license enables third parties such as Google to reuse structured information from Wikipedia and/or to use the platform as a host for their own structured data.[3] Now, a number of websites and applications embed data extracted via Wikidata, and some external data repositories, most notably Google's Freebase, have been discontinued and their data and API migrated to Wikidata. This has led to a growing centralisation of factual data housed in Wikimedia projects.
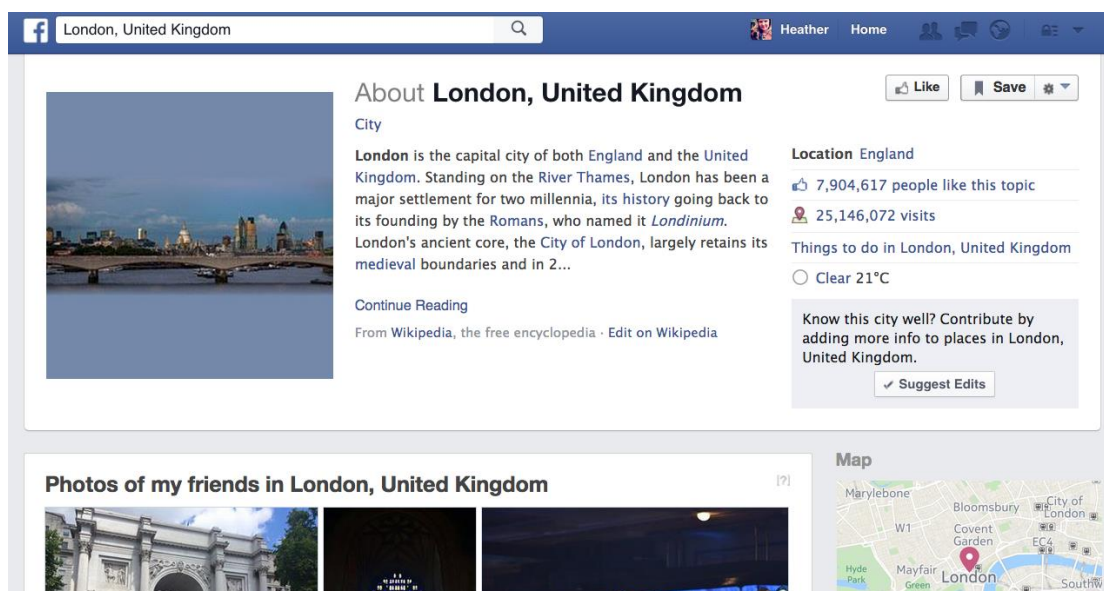
Figure 1.2 Screenshot of search results on Facebook for 'London' displaying information extracted from Wikipedia

Source: Facebook.com, 11 June 2015

As these changes were implemented relatively recently at the time of this study, and were also largely unexplored in terms of their impacts on the perception of facts online, this study has been designed to take advantage of this. Facts have enormous power to

---

[3] Data is structured when it resides in fixed fields according to a data model that standardizes how different data elements relate to one another.

determine who the winners and losers are in important political battles. Some of the more notable battles over facts have taken place over the existence of climate change and the dangers of tobacco smoking (Oreskes & Conway, 2011), but battles about how authoritative statements are structured are a feature of daily life – from the ways in which statements about our health are made by medical practitioners to the naming of political groups like the group that calls itself the Islamic State (Dathan, 2015).

Because of this power, the platforms that are more trusted than others to display facts about the world have become important sites of struggle in which different groups of actors vie for control over their representation. As Morgan (2010b, p. 4) notes, certain facts travel better than others, and it is 'the possibilities for facts to travel well that is important to our lives.' In countries where facts about HIV-AIDS have not travelled well because they have been deemed to be illegitimate, for example, the dangers of the epidemic have been exacerbated (Morgan, 2010b, p. 4).

These struggles are increasingly coming to the attention of the news media as they start to investigate the representation of phenomena that are important features of public debates. Recently, journalists have covered the speed at which English Wikipedia changed the article on Bruce Jenner[4] to represent her new name and gender; one journalist wrote that this signified the 'urgency of updating an information source that everybody uses' (Ramos, 2015). Wikipedia administrators' management of the gamergate[5] controversy was also recently subject to review when the Guardian

---

[4] Caitlyn Jenner (was Bruce Jenner) is an American athlete who came out as a transgender woman in April 2015.

[5] The gamergate controversy occurred when several women from the video game industry were subjected to a campaign of misogynist attacks from August 2014.

published a story about Wikipedia's arbitration committee voting to ban certain editors from editing articles about feminism (Hern, 2015).

Not only are the news media starting to cover how Wikipedia represents the world, but also how politicians, public officials and corporations are editing articles in which they have a potential conflict of interest. In March 2015, a number of media outlets in the United States covered how the New York Police Department was caught editing articles on police brutality (Mathis-Lilley, 2015) and during the 2015 British election campaign, there was a controversy around an IP address linked to Tory MP, Grant Schapps being banned from editing (Ramesh, 2015). The Wikimedia Foundation strongly discourages editors who have a conflict of interest from editing articles directly, and banned hundreds of accounts from a public relations firm in 2013 for making edits that were determined to be subverting Wikipedia policies (Arthur, 2013).

Wikipedia's portrayal of phenomena has become so influential that many governments recognise it as a critical platform for propaganda. Journalists are regularly tipped off by Wikipedia editors or by automated engines that publish details of articles relating to local politics edited by users under government IP addresses. In one example, editors from inside the Kremlin and Kremlin supporters tried to change an article on Russian Wikipedia relating to the MH17 disaster to reflect the narrative advocated by the Russian government that a Ukrainian jet had tried to shoot down the plane. In the United States, the entire US Congress IP address range was banned after repeated vandalism, among which was an edit claiming that Donald Rumsfeld was 'an alien lizard who eats Mexican babies' (Miller, 2014).

Scientists and academics are also starting to recognise the importance of having their research reflected in Wikipedia articles as an effective method of amplifying their ability to communicate research results to the broader public (Teplitskiy, Lu, & Duede, 2015). With the increased pressure on scientists and academics to indicate the public impact of their work, researchers are becoming increasingly interested in whether Wikipedia articles contain citations to their work (Reich, 2011) and in promoting themselves in articles about their work (Elvebakk, 2008). Academics also use Wikipedia for their own research (Weller, Dornstädter, Freimanis, Klein, & Perez, 2010), either as a source of background reading or in citations (Dooley, 2010).[6]

Although some academics have been able to successfully edit Wikipedia articles with citations to their research, there have been some notable failures. In 2012, history professor Timothy Messer-Kruse wrote about his experience trying (and failing) to edit the article about the Haymarket riot, a subject that he had been researching for 10 years (Messer-Kruse, 2012). Editors of the article continuously reverted Messer-Kruse's edits when he wrote that earlier analysis of the riots was incorrect. Editors declared firstly that his edits constituted original research and needed to be attributed to a reliable source. When Messer-Kruse published an academic book about the subject years later and cited the publication, his edits were once again reverted because editors declared that he had a conflict of interest.

It isn't only in the areas of history and biography that conflict occurs. A list of controversial subject areas on Wikipedia includes science, economics, linguistics and the environment, along with politics, religion and people. The list of controversies regarding

---

[6] Citations to Wikipedia are still controversial in many fields where a citation to Wikipedia would be frowned upon.

science, biology and health include topics relating to AIDS, aspartame, intelligent design and obesity, thus indicating the ways in which Wikipedia mirrors debates in larger society.[7] Wikipedia is, however, not a perfect mirror of knowledge – no representation is. It is therefore important to understand how such representation occurs and who (or what) has the power and authority to determine what Wikipedia represents. Wikipedia seems to be both signalling the rise of new centres of expertise and authority in those who are able to effectively edit Wikipedia while at the same time as reaffirming certain knowledge authorities in its choice of sources. Analysing exactly how authority on Wikipedia is constructed empirically is the goal of this thesis.

1.2   The demise of the gatekeeper and the rise of the amateur

The dominant theory surrounding Wikipedia and other participatory production systems including free and open source software, citizen science, citizen journalism and volunteered geographic information is founded on the idea that the Internet has enabled the removal of the gatekeeper figure and the rise of the everywo/man. This idea is present in both scholarship and media discourse, and is highlighted by the use of imagery where power is being wrested from traditional publishing institutions, institutions that house knowledge workers, as well as institutions supporting the education and training of knowledge workers. The Internet and its free and open principles meant that no longer did people require traditional academic and media publishers to distribute their writing and creative work, no longer were institutions required in order to gain access to expensive equipment, no longer did people require educational institutions in order to be educated or certified. When information was free

---

[7] See https://en.wikipedia.org/wiki/Wikipedia:List_of_controversial_issues.

and accessible, institutions would crumble and with them the gatekeepers who prevented the free flow of information.

The 2006 Time Magazine Person of the Year Award perhaps best illustrates this zeitgeist. In December 2006, Time Magazine declared that its person of the year was 'you'. The cover of the magazine featured a computer screen with the words, 'You. Yes, you. You control the Information Age. Welcome to your world' (Grossman, 2006). The editorial argued that ordinary people now controlled the means of producing information and media because they dissolved the power of the gatekeepers who had previously controlled the public's access to information.

> [2006 is] a story about community and collaboration on a scale never seen before. It's about the cosmic compendium of knowledge Wikipedia and the million-channel people's network YouTube and the online metropolis MySpace. It's about the many wresting power from the few and helping one another for nothing and how that will not only change the world, but also change the way the world changes. (Grossman, 2006)

The idea that the 'many' were 'wresting power from the few' was shared by a host of commentators and scholars at the time. This democratic ideal was inspired by the belief that many more people were now doing work that had previously been done by credentialed individuals within large institutions. Now, non-academics could write encyclopaedia articles, laypeople could produce films, and concerned citizens could produce news articles. With the decrease in the costs of the means of production and the connection of millions of ordinary people to a network of potential audience members, co-producers, employees and publishers, anyone could be a journalist, an engineer or a scientist (Gillmor, 2008; Leadbeater & Miller, 2004; Shirky, 2009); now anyone could be an expert in something (Weinberger, 2011). We had moved from a 'read only' culture to a 'read write' culture (Lessig, 2009) that was characterised by active cultures of participation (Jenkins, 2006) instead of passive consumption.

Yochai Benkler (2006) offers one of the prevailing theories for explaining Wikipedia as a platform for the free expression of people who are unencumbered by the gatekeepers of the past. Benkler argues that there have been significant changes in the organization of information production that have resulted in the rise of nonmarket and non-proprietary production, where individuals are able to take a more active role than was previously possible in the production of information goods. This rise of individual agency leads to an inevitable clash with previous hierarchical market-driven industries which will inevitably decide the fate of each of these models. Benkler uses Wikipedia as a key example of nonmarket, non-proprietary and non-hierarchical 'peer production' where incentives to create cultural goods are based not on price signals but by pro-social goals. He extends the idea of 'pro-social goals' in a paper with Helen Nissenbaum (Benkler & Nissenbaum, 2006) which asserts that peer production offers opportunities for people to exhibit and experience virtuous behaviour, and enables positive character formation among those who participate in projects like Wikipedia.

Benkler's framing of peer production is based on two key theories of power relating firstly to the power of individuals who co-create peer production products like Wikipedia, and secondly to the products that are developed through peer production. Firstly, Benkler argues that individuals' autonomy is on the rise. Because it is cheap to communicate in the networked public sphere, individuals can represent their own interests as well as loosely associate with others with similar interests, and thereby avoid the gatekeeping power of the media.

> The various formats of the networked public sphere provide anyone with an outlet to speak, to inquire, to investigate, without need to access the resources of a major media organization. (Benkler, 2006, p. 11)

Because individuals are independent from organisations, corporations, and the market, the assumption is that individuals have equal power to speak because they can very easily (that is, cheaply) access the means of producing information. Peer production products such as Wikipedia are able to avoid bias, according to Benkler, because corporate media pressures are not there to taint their editors. Benkler argues that the new 'networked communication environment enables many more people (to) connect their perspectives to many others and to do so in a way that cannot be controlled by media owners and is not as easily corruptible by money as were the mass media' (Benkler, 2006, p. 11).

Secondly, Benkler writes that information products in the networked public sphere can never have too much power (unlike their corporate media counterparts) because there are many alternative products to choose from, and it is difficult to buy attention or use money to squelch an opposing view. In other words, the networked public sphere enables parallel alternative visions of the world where individuals are not reliant or dependent on the mass media. Benkler compares the power of individuals and information products within the non-market sphere with that of corporate media, which he argues restricts particular points of view because there are too few gatekeepers in relation to the interests that need to be served, and because corporate media tends to concentrate on politically unengaged programming in order to gain the most profit.

In each of these areas, Benkler compares non-market peer production favourably to corporate media. Whereas corporate media suffer from too few gatekeepers to represent diverse interests, individuals can represent their own interests in the networked information sphere; whereas corporate media owners have too much power to shape opinion, diverse groups of individuals collaborating together produce a range

of alternatives; whereas corporate media produces politically unengaged programming, the networked public sphere enables individuals who are unhindered by these restrictions to produce politically engaged, reflexive programming, thereby performing the watchdog function the media was previously tasked with.

The image of the destruction of the gatekeeper, in the form of either the knowledge institution or the knowledge worker has been emphasised by other scholars. David Weinberger (2011), for example, writes that the Internet has enabled the destruction of the gatekeeper figure who took the form of the editor or curator.

> No editors and curators who get to decide what is in or out. No agreed-upon walls to let us know that knowledge begins here, while outside uncertainty reigns – at least none that everyone accepts… The Internet is what you get when everyone is a curator and everything is linked. (Weinberger, 2011, p. 45)

Weinberger's belief is that, in the past, '(e)xperts were a special class' (Weinberger, 2011, p. 67) where in order to publish books, people had to pass through editorial filters, but that 'On the Net, everyone is potentially an expert in something' (Weinberger, 2011, p. 67).

Accompanying the theme of the gatekeeper's demise was an image of the rise of a different figure, that of the amateur (Keen, 2008; Leadbeater & Miller, 2004; Shirky, 2009). Amateur identity is defined by the independence of an individual from the institutions and organisations that previously housed expert identities, and is signified by non-professionals becoming involved in the production of news ('citizen journalism'), science ('citizen science') and engineering through free and open source software projects. Previously, the work of journalism, science and engineering had been confined to those working within (and certified by) institutions that employed them as journalists, scientists and engineers. The rapidly diminishing cost of the means of producing information led to disruption, not only in the ways in which information was

being accessed, but also in the identities of those who were said to be producing information.

For Dan Gillmor (2008), a new cadre of citizen journalists (previously the audience) was now producing news outside the purview of the few large media conglomerates. Clay Shirky (2009) noted that the wide availability of tools for organizing and communicating has led to the 'mass amateurization of society' that breaks previous definitions of journalism, journalists and journalistic privilege. This is the age, or the 'cult' according to Andrew Keen (2007), of the amateur.

A number of scholars have responded to Benkler, Shirky and others who have heralded participatory media as liberatory. Nathaniel Tkacz (2012), for example, argues that 'discourses of collaboration are, like openness, depoliticized' (p. 82) by those who tend to downplay the organising forces within collaborative work but that there is an invisible politics at work in projects like Wikipedia. Kress, Finn and Turner (2011) doubt peer production's revolutionary potential and note that there are consequences of 'peer production's failure to develop institutional mechanisms that secure bureaucratic values such as inclusion, explicit rule-making, accountability and institution persistence' (p. 14). Others have developed empirical studies in order to engage with theories of peer production projects. Studying the Wreckamovie peer production film community, Isis Hjorth (2013) challenged core assumptions associated with peer production of culture, particular in relation to the supposed distinction between networked cultural production and peer production. In the realm of empirical research on Wikipedia, there are similarly conflicting reports from the field.

1.3   Stratification and skew

Despite the chorus of voices declaring this to be an age of active participation in the production of knowledge and culture, there is evidence of a significant process of social stratification along a number of different lines within peer production communities where participation is, according to the rhetoric, open to everyone. Whereas Facebook and Twitter have more or less equal numbers of female and male users (Duggan, Ellison, Lampe, Lenhart, & Madden, 2015), demographic studies of Wikipedians (Glott & Ghosh, 2010; Hill & Shaw, 2013; Lam et al., 2011), citizen scientists (TTFNROB, 2015) and map-makers (Stephens, 2013) indicate that there is a significant gender, geographic and socio-economic skew in who participates in open projects. Demographic studies of Wikipedians indicate the highest skews such that between 84 and 90% of Wikipedia editors are men (Glott & Ghosh, 2010; Hill & Shaw, 2013; Lam et al., 2011), the majority have tertiary education and a significant number of editors across language versions speak English (Wikimedia Foundation, 2011).

Some scholars argue that Wikipedia's representations of the world reproduce existing asymmetries. Mapping geotagged articles in English Wikipedia, Mark Graham finds that 'almost all of Africa is poorly represented in Wikipedia' (Graham, 2011, p. 275). On the issue of Wikipedia's gender skew, Lam et al. (2011) finds that the low proportion of females participating in English Wikipedia resulted in measurable imbalances relating to content quality so that articles relevant to women are significantly shorter and have lower assessment ratings than those interesting to men (Lam et al., 2011, p. 6).

Similarly, Reagle and Rhue (2011) find that, although Wikipedia biographies on women are longer and more numerous than Encyclopaedia Britannica in absolute terms, 'Wikipedia articles on women are more likely to be missing than are articles on men relative to Britannica' (p. 1138). Analysing accounts of Singaporean and Philippine

27

history on Wikipedia, Luyt (2011) argues that, despite the potential of new media for making visible previously marginalized voices, a more likely outcome is a reproduction of the status quo in historical representation.

Other studies have indicated that Wikipedia's representations do not only mirror existing asymmetries but can actually exacerbate those asymmetries. Graham, Hogan, Straumann, & Medhat (2014) argue that it isn't only issues of connectivity that prevent people in developing countries from contributing to Wikipedia and that representation is a vicious cycle for those with strong editing cultures in local languages, while those on the peripheries of these countries fail to reach critical mass (Graham et al., 2014, p. 14). Similarly, Joseph Reagle (2013) extends these findings with a qualitative study in which he argues that low female participation in free culture communities, particularly within Wikipedia, is the product of a culture that is alienating towards women and that gender disparities are even worse in Wikipedia than in the computing culture from which it arose. Reasons for this include that geek stereotypes can be alienating, open communities are especially susceptible to difficult people, and the ideas of freedom and openness can be used to dismiss concerns and rationalize the gender gap as a matter of preference and choice.

Other scholars have investigated the process by which Wikipedia editors decide what should be included or excluded from the encyclopaedia. A number of articles are deleted every day on Wikipedia, either according to a process of deliberation amongst editors or a process in which an administrator unilaterally decides to delete an article according to criteria for speedy deletion. A study of article deletions on the English Wikipedia that I undertook with Stuart Geiger in 2011, for example, showed a clear division among experienced and inexperienced users. The article deletion process is managed by a

relatively small number of longstanding users, and the majority of deleted articles are deleted under the criteria of 'no indication of importance' rather than for spam, copyright violations or *patent nonsense* which constituted only about 6% of all deletions. We also found that the majority of deletion discussions have very few participants, most of whom have previously participated in such discussions as experienced users (Geiger & Ford, 2011).

This growing stratification of users along the lines of experience is extended in work by Schneider, Samp, Passant & Decker (2013). The authors analysed argumentation patterns that are used in evaluation and decision-making on Wikipedia using a sample of the 500 deletion discussions (on average) that take place every week on Wikipedia. They found that familiarity with Wikipedia's policies and norms correlates with newcomers' ability to craft persuasive arguments, and that acceptable arguments employ community-appropriate rhetoric that demonstrates knowledge of policies and values. There are 56 English Wikipedia policies, about a hundred guidelines and hundreds of essays about Wikipedia norms and values. Knowing how to speak in Wikipedia's complex language seems to be increasingly important to being a productive member of the community.

1.4   The rise of new expertise?

Expert identity is not only the result of an individual's technical mastery over a particular subject, independent from her social context. Numerous scholars have recognised that identity is, in fact, as much a social phenomenon as it is an outcome of technical prowess (Goffman, 1959; Jasanoff, 2004; Maasen & Weingart, 2006). A person can call herself a journalist on her blog, for example, but she may very well not be able to claim journalistic privilege if the state calls on her to release information about her

sources. There is no intrinsic quality to any particular representation of knowledge; what is important is understanding why certain knowledge claims are successful and others are unsuccessful (Bloor, 1984).

Expertise is socially constructed and is not directly related to any intrinsic quality of an idea, but rather to whether society (in the form of institutions such as academia or the state) gives credence to what one says or does. Ordinary people may now be performing work previously done by professionals, but it doesn't follow that they have all been accorded the same credibility or the same power as the gatekeepers they are supposedly supplanting.

The narrative about ordinary people taking over from the mass media assumes a clean break between one social structure and another, but historically technologies never produce such radical change. Societies still show dependencies on old technologies and traditional power centres long after they become unnecessary (Edgerton, 2008). Putting the question of expertise within its historical context provides a much more complex picture about the origins of the disruptions that we are currently witnessing. A more contextual, historical approach to the formation of identities is necessary in order to understand which groups are in control of knowledge representation in the age of the network.

According to Maasen and Weingart (2006), debates about the nature of expertise are not new, but are re-surfacing because of three key parameters that have changed in society. Firstly, industrialised countries have become increasingly democratized with a growing proliferation of organisations operating outside the sphere of formal political parties. Secondly, the 1960s anti-nuclear and environmental debates saw scientists being drawn in to represent both poles of a debate, with the consequence that the public saw science

as not presenting a single view, and that scientific knowledge could be contradictory, incomplete and biased. This resulted in a decrease in the authority accorded to scientific knowledge, and the conclusion that scientific knowledge could no longer be taken as neutral, objective and reliable. Finally, the democratization of expertise has seen the demystification of scientific knowledge and of scientists, so that now scientific knowledge is seen as uncertain, risky and incomplete (Maasen & Weingart, 2006, p. 2).

These three key changes have led to a situation where, in spite of a loss of authority of the traditional scientific expert, the reliance of policy makers on experts to help craft policy regarding an ever-expanding number of niche topics has increased significantly (Maasen & Weingart, 2006, p. 4). Technology has certainly helped to enable access to expertise beyond the academy and to enable experts to widely disseminate their findings, but Maasen and Weingart make an important point about the need for historically situated accounts of expertise that account for the ways in which there is stasis as well as change in the way that notions of expertise circulate in society.

Sheila Jasanoff (2004) offers the theory of co-production for conceptualising changes in technology and their impact on expert identities. According to Jasanoff, at moments of significant social and technological change, there is a re-ordering of social structures often by reaffirming the legitimacy of existing social arrangements. During times of change, a host of new actors begin to participate in the production and distribution of knowledge in order to fill gaps in expertise that result from the adoption of new technologies. In response to change, societies must redraft rules of social order by reaffirming some roles and introducing others.

The theory of co-production is developed out of the field of Science and Technology Studies (STS) which problematizes the role of science and technology in social change by laying bare the socio-technical processes by which science and technologies advance, and by interrogating how the social and the technical are imbricated with one another during such change. STS scholars investigate the social and technical contexts in which scientific knowledge and technologies are produced, asking questions about why certain knowledges are accepted as true, scientific facts, and others as myths, opinions or ideologies.

In order to understand how knowledge is taken up and who is involved in its construction and performance, a fruitful method is to study the ideological discourse that actors engage in as they construct and debate competing claims about the world (Geertz, 1973; Gieryn, 1983, 2001). Discourse does not only involve words, however, but also the enactment of conventions. We do 'things with words' but we also do 'words with things' (Latour, 1991). Digital speech acts (Isin & Ruppert, 2015) are thus a critical component of understanding how representations of knowledge are both constructed and performed in the online environment.

This discursive process by which knowledge claims are constructed and debated is neither purely social nor purely technological. Wikipedia's socio-technical environment, for example, is heavily mediated by code and code plays a significant role in determining what Wikipedia represents. Some scholars have started to investigate the effects of Wikipedia's coded environment in the area of bots, the automated actants developed by Wikipedia editors in order to perform automated editing tasks on the encyclopaedia.

Geiger and Ribes (Geiger & Ribes, 2010), for example, demonstrate the role of non-human actors in the process of vandal banning, arguing that the decentralized activity enabled by automated and semi-automated tools on Wikipedia is a type of 'distributed cognition' (p. 117). Geiger (2011) extends this work to show how bots are exercising control over a vision of what the encyclopaedia should be and how editors should work together within it. In describing the increasing role of algorithmic actors on Wikipedia, Geiger demonstrates the ways in which bots produce order and enforce rules on the encyclopaedia.

Similarly, Niederer and van Dijck (2010) investigate the increasingly important role of bots in the rise of Wikipedia, arguing that it is impossible to understand Wikipedia's response to vandalism without an appreciation of the encyclopaedia as a sociotechnical system driven by collaboration between users and bots. The authors provide a schematic overview of Wikipedia users according to their permission levels, and note that the permission level of bots is below that of administrators but well above the authority of registered users (Niederer & van Dijck, 2010, p. 1373). Wikipedia's content management system, argue Niederer and van Dijck, allows for 'protocological control, a mode of control that is at once social and technological – one cannot exist without the other' (p. 1373).

Power in these studies is viewed as a product mainly of technical permissions. Unregistered users, for example, do not have permission to edit protected pages, and administrators have the ability to ban users and award permissions to other users. Arazy, Nov and Ortega (2014) attempt to reveal Wikipedia's organisational hierarchy by studying the access privileges of almost 5 million Wikipedia members of which about 10,500 hold special access privileges. The authors identify a set of privileges that are

granted to users with particular roles, including the ability to grant (or remove) the access privileges of other contributors (Arazy, Nov & Ortega, 2014, p. 10). The result is a series of six levels of power from the benevolent dictator at the top of the hierarchy to ordinary users at the bottom (see figure 1.3 below). According to the authors, this hierarchy represents an ecology of roles necessary for quality assurance, coordination, and conflict resolution in online communities (p. 12).
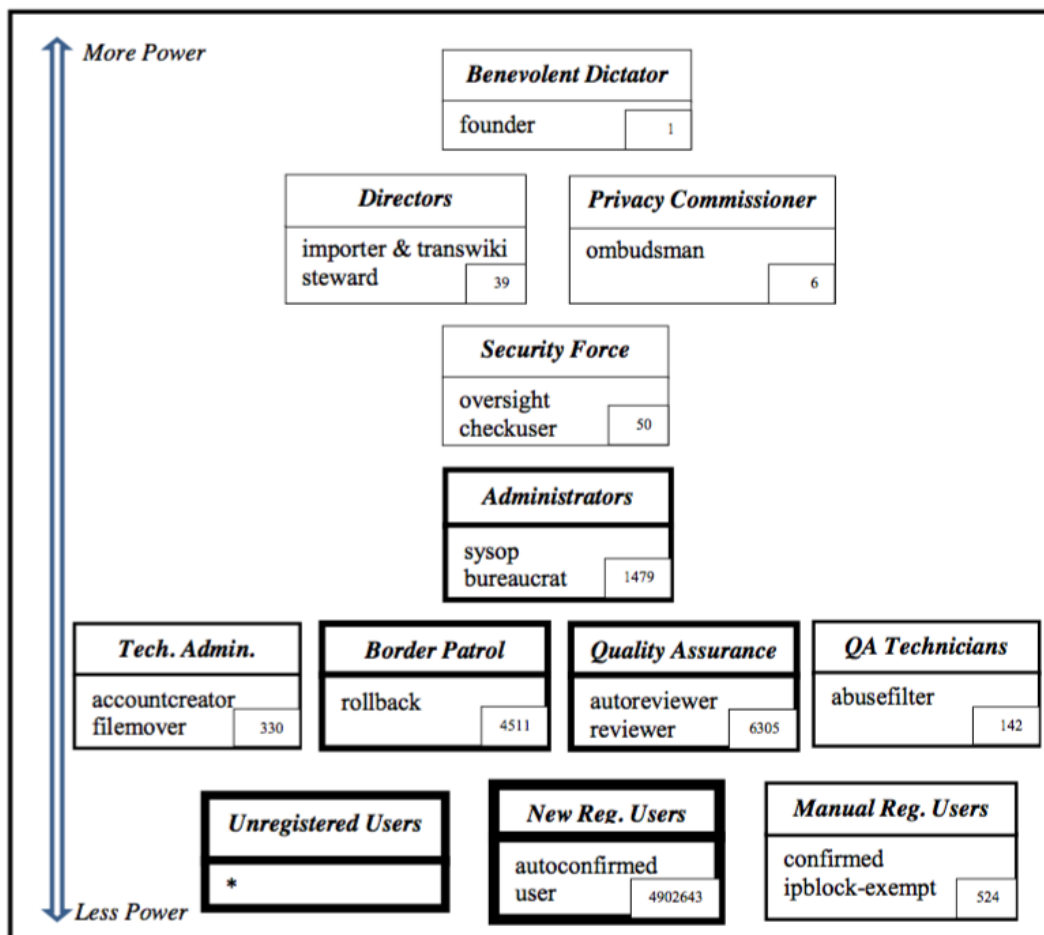
Figure 1.3 Wikipedia's organizational chart according to Arazy, Nov & Ortega (2014)

Note: Thickness of borders corresponds to the number of participants performing the role. Note that it is not possible to determine the number of unregistered participants, since an IP address cannot be linked to a single user.

Source: Arazy, Nov & Ortega, 'The (Wikipedia) world is not flat: On the organizational structure of online production communities' (2014)

The power to have one's edits prevail on Wikipedia, however, is not solely a function of editing permissions. Contributors with the same or similar levels of technical permissions are often in debate with one another about how phenomena should be represented, and even administrators, who have some of the highest levels of permissions, do not always use the power granted to them without following processes dictated by policy and norms.

Another group of studies has focused on the role of social forces in guiding decision making on Wikipedia with a particular focus on the role of policy. Policies are an important element of social interaction on Wikipedia (Bryant, Forte, & Bruckman, 2005; Pentzold & Seidenglanz, 2006; Viegas, Wattenberg, Kriss, & van Ham, 2007). Analysts have found that policies are used to appeal to authority in order to justify a contributor's changes to an article, and that policies provide a common resource for new users to learn about editing and behavioural conventions (Viegas et al., 2007). There has been a significant rise in the number of policy and other administration pages on Wikipedia: between mid-2003 and late 2005, the number of administrative pages grew at a rate of nearly eight times that of main article pages (Viegas et al., 2007). Wikipedia policies are numerous and complex, and encompass so many levels of authority that a user's relatively greater understanding of policy enables them to more effectively participate in debates in order to influence representation (Ford & Geiger, 2012).

A study of talk pages by Pentzold and Seidenglanz (2006) relies on the writings of Foucault to argue that Wikipedia policy is an important feature in defining the rules by which participants 'delimit the sayable, define legitimate perspectives and fix the norms for the elaboration of concepts' (p. 65). In response, Kriplean, Beschastnikh, McDonald, & Golder (2007) argue that Pentzold and Seidenglanz overemphasize the power of

policy in defining what editors can do on Wikipedia. Instead, they argue that, although policy enables collaboration 'by providing a common language and strategies for action that contributors can draw on to interpret and apply to difficult or unanticipated situations' (p. 171) policies don't translate into obvious actions, and their ambiguity leads to 'power play' (p. 172) rather than already defined outcomes.

Power play is defined by the authors as 'an attempt by an individual or a group to claim legitimate control over an article' (p. 172) and the authors develop a list of seven forms of power play from their sample of talk pages. These include the delimitation of an article's scope by certain contributors ('article scope'), the presentation of decisions made in the past as absolute and uncontested ('prior consensus'), the bolstering or undermining of a position based on the legitimacy of a contributor in terms of their expertise ('legitimacy of contributor'), and the discrediting of a source ('legitimacy of source') (p. 172).

Kriplean et al. argue that their approach is based on Giddens' criticism of Foucault where, instead of overemphasizing the power of the institution and neglecting the actions of individuals, they focus, instead, on how 'groups of contributors claim legitimate control over content through the discourse of policy' (p. 170). These two approaches are not, however, too distinct from one another. Whereas Pentzold and Seidenglanz focus on the role of policy in determining what can be said on Wikipedia, Kriplean et al. investigate how language is used strategically in order to retain power by some over others.

One of the key stakeholders in Wikipedia that has had less attention by scholars are the authors of scientific reports, the academics writing books about subjects covered by

Wikipedia, as well as the journalists, bloggers, website builders and social media authors who constitute Wikipedia's sources and citations. Since policy dictates that all Wikipedia articles must be based on 'reliable sources' (Wikipedia:Verifiability)[8] the availability of sources, the perspectives that those sources represent and how they are evaluated by editors are all important factors in deciding what facts Wikipedia represents and which facts it ignores.

Empirical research in the analysis of Wikipedia sources has so far analysed the types of sources that are represented on Wikipedia (Ford, Sen, Musicant & Miller, 2013; Luyt, 2011, 2012; Luyt & Tan, 2010; Nielsen, 2007) as well as how source use relates to the geographic region in which contributors are located (Sen et al., 2015). Scholars have found that certain types of sources are preferred by Wikipedia editors and that this has an impact on how Wikipedia covers different topics. Luyt and Tan (2010), for example, found that Wikipedians' preference for United States government sources and online sources represents a particular point of view about world history, and Luyt (2012) found that Wikipedians' preference for short texts that can be more easily mined means that Wikipedia articles suffer from the 'congealed consensus of the institutions hosting them' (p. 1873). Luyt argues, in conclusion, that Wikipedia source preference is based on the assumption 'that texts are undifferentiated bearers of extractable facts' (p. 1876). As a consequence, Wikipedia articles tend to include 'long reams of facts rather than persuasive analysis and interpretations' (Rosenzweig, 2006).

Other studies have investigated the role of source conventions and norms in the representation of facts. Digital artists, Scott Kildall and Nathaniel Stern (2011), provide a

---

[8] URLs and other citation information for Wikipedia documents are provided in the 'primary sources' section of the references.

notable example in discussing their artwork entitled 'Wikipedia Art' which they call a 'collaborative performance and a public intervention' (Kildall & Stern, n.d.). Kildall and Stern created a kind of future fact on Wikipedia's platform by writing an article about the work and citing sources that they simultaneously created in order to reference the work. Kildall and Stern used the piece to argue the authoritative nature of Wikipedia, where something 'becomes true when on Wikipedia' when citations are used as a 'performative act' (Kildall & Stern, 2011, p. 174).

Both the technical and the social studies of power and decision-making on Wikipedia offer useful insights into the process by which decisions are made, but there are two clear gaps in the way that power is defined and analysed within these studies.

The first is that there is a clear division between scholars' conceptions of power on Wikipedia as either technologically determined (predominantly through technical permissions or the actions of automated agents) or determined via social mechanisms (predominantly relating to policies and norms). Methodologically, this means that studies involve analysis of either talk pages or edits but not both of these together. These approaches tend to treat technical agents and contributors separately, without accounting for the ways in which the social and the technical are imbricated with one another. On Wikipedia, symbolic action takes numerous forms, only one of which is the type of talk that takes place on talk pages of article.

Secondly, studies tend not to focus on the subjects of the article themselves, focusing rather on the process by which editors make decisions within the context of Wikipedia's socio-technical structure. The subjects of the article about the city of Johannesburg or the victims of the latest natural disaster or the sufferers of HIV-AIDS are important

stakeholders in the flow of facts through Wikipedia's socio-technical system. Wikipedia is an encyclopaedia made up of facts that are becoming authoritative for a wide range of stakeholders. Understanding the impact of the representation of facts on the subjects of those facts is therefore as important as understanding who participates in Wikipedia's community.

This thesis responds to each of these gaps by analysing the factors that play a role in the travel of facts (following Morgan, 2010, as touched on earlier and explained in more detail in Chapter 2) through Wikipedia's socio-technical system. In doing so, we come to understand how Wikipedia is reaffirming certain spokespersons of facts (in the form of sources and citations) while relying on particular types of expertise (in the form of Wikipedia editors) as well as who is left out of this process.

Instead of dealing with technological and social forces separately, I analyse how the technological and the social are imbricated with one another in complex ways in the form of policies, tools, norms and discourses. Instead of analysing only talk pages or edit logs, I analyse all actions taken by actors and actants as the article progresses over time, including discussions taking place outside of the talk page on Wikipedia's meta pages that are used to discuss issues of common concern, as well as on the multiple mailing lists, blogs and media reports in which discussions about articles take place. Instead of only focusing on editors, I also account for actors such as sources (and their authors and publishers) as well as the subjects themselves, who may or may not have a voice in making decisions over the article.

1.5   Research design and objectives

This thesis considers Wikipedia as a socio-technical system composed of multiple intersecting networks composed of both social relations between people and technical aspects relating to organisational structure and processes. The term, socio-technical system was derived originally from the field of organisational science, particularly research by Trist, Bamforth and Emery on workers in English coal mines at the Tavistock Institute in post-World War II Britain.[9] Wikipedia's socio-technical system is comprised of Wikipedia's complex infrastructures and the human behaviour that accompanies those infrastructures. Infrastructure, here, is defined according to definitions posited by infrastructure studies scholars who take a broad relational view of infrastructure that extends beyond bricks and mortar, tubes and wires, to encompass the wide range of institutional, social practices, norms, standards and other artifacts (Jackson, Edwards, Bowker, & Knobel, 2007; Sandvig, 2013; Star, 1999). Information infrastructure, in other words, consists of the often-invisible network of objects, tools, standards and practices that govern the way that we use the internet (Sandvig, 2013).

Wikipedia can be seen as consisting of infrastructures that maintain its functioning; it can also be seen as infrastructure for the Web itself. Wikipedia is comprised of a large network of artifacts, norms and practices that both support large-scale distributed work and reflect people, places, events, and concepts from the world around us. Wikipedia infrastructures consist of infrastructures for citation management, for vandalism prevention, and for the distribution of tasks, amongst others. It is also, in itself, an infrastructure for other knowledge communities who use the data extracted from

---

[9] See, for example, Trist & Bamforth (1951).

Wikipedia to seed their own knowledge bases (for example, search engines and other apps as indicated earlier in this chapter).

In designing this research project, particular attention was paid to exposing the often-invisible network of objects and processes, tools and practices, conventions and vocabulary that are critical to Wikipedia's daily functioning. Wikipedia is just one of sixteen projects supported by the Wikimedia Foundation (see figure 1.4). Rather than being bound by a particular project or language version[10], this study followed the progress of sets of facts as they progressed through the network of sites housed by Wikimedia and beyond those sites as facts find new users and uses.

Facts in Wikipedia are represented in the title of articles, the categories inserted into articles, the series of facts that make up the body of articles and the links that cross-reference the article with other articles inside and outside of Wikipedia. Furthermore, facts travel beyond their initial publication on Wikipedia to other language versions, to other Wikimedia projects, and outside of Wikipedia as facts are extracted for use on external websites and applications. Conventions that originate in Wikipedia are also taken up by other organisations and individuals but have origins in earlier epistemologies and discourses (as explained in chapter five).

The following of facts in this thesis, then, involved the analysis of representations within a variety of Wikimedia's separate sites: the images and audio files embedded in Wikipedia articles that were originally published on Wikimedia Commons, discussions on mailing lists administered by the Wikimedia Foundation that referred to project

---

[10] Although numerous versions of facts represented by Wikipedia's multiple language versions were followed, English Wikipedia was privileged in this study.

documents on Meta-Wiki, discursive structures embedded within Wikimedia Foundation documents, and Wikidata's structured representations of facts that are embedded within numerous language editions of Wikipedia.
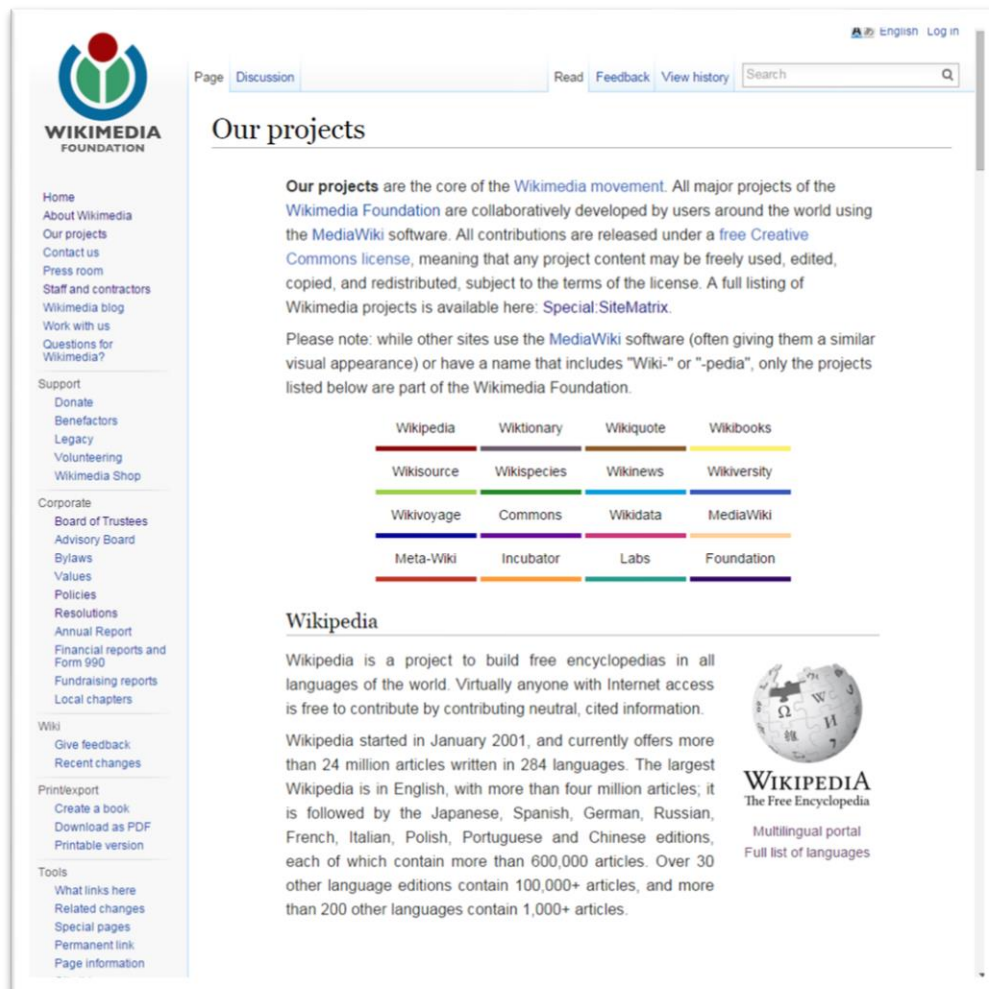
Figure 1.4 Wikimedia Foundation projects

Source: https://wikimediafoundation.org/wiki/Our_projects

The two sets of research questions that this thesis answers are defined according to the practical tracing of facts as they traverse the networks of actors and practices that become enrolled after facts are first published on Wikipedia, as well as the impact of the

dynamics of such travel in the sources of authority and expertise. Research questions are thus defined as follows:

> RQ1: Can facts be shown to travel within Wikipedia's socio-technical system? If so, do some facts travel further and with greater integrity than others, and why is this so?

> RQ2: Is there evidence that the roles of knowledge authorities and experts are being reconfigured within Wikipedia's socio-technical system, and if so, how has this reconfiguration taken place?

Analysing how facts travel within Wikipedia's socio-technical system at a granular level provides evidence that demonstrates how certain authorities are more acceptable than others to Wikipedians and how particular expertise is required in order to enable the travel of facts within Wikipedia. Wikipedia constitutes a filter of facts that is both socially and technologically constituted. Understanding the mechanics of that filter is the primary objective of this thesis.

In order to implement this objective, I have employed the methods and principles of ethnography. The field site has been constructed as a network (Burrell, 2009) by following the progress of a series of articles over time, and by defining the human actors, the non-human actants or objects, as well as the processes and practices that play a role in how the subject is represented. Once the network was established, I investigated each of the actors and actants through interviews and further following of related documents, and initiated series of participant observation activities in order to understand the mechanics of related practices.

Three networks or followings were studied in the collection of data for this project. The first involved following the reliable sources policy on Wikipedia as a way of understanding the terrain in which facts travel. The second and third networks were constructed by following the progress of two groups of facts through Wikipedia's socio-technical structure. Surr is a sport played in rural northern India and the article representing it faced enormous challenges, stopping relatively short in its travels through Wikipedia and to related databases. The 2011 Egyptian Revolution, on the other hand, travelled rapidly and widely to a number of different language versions and produced numerous related facts/articles.

In both cases, surr and the Egyptian Revolution are representative of more than just the subjects they portray. Surr represents an attempt to extend the boundaries of Wikipedia to cover issues about which there are little (if any) reliable, published sources according to Wikipedia's definitions. The Egyptian Revolution represents a continued expansion of the boundaries of Wikipedia to include facts relating to breaking news. Both facts have a political impact on their subjects. The acceptance or rejection of a cultural activity in rural India is symbolic of whether colonial attitudes towards former colonies have been abandoned or whether those attitudes have merely taken on a new form. The representation of popular protests in Egypt's Tahrir Square has an impact on the legitimacy of future governments. While this is not a study designed to argue that these cases are representative (in a probabilistic sense) of what happens in Wikipedia, I show how they are emblematic of a particular range of ways that facts can, and as I will show, do travel.

1.6 Conclusion

Every day, thousands of decisions are being made on Wikipedia about which facts should be included and how they should be represented. These decisions have an impact on the visibility (or invisibility) of subjects not only on Wikipedia's platform but on the sites and databases that extract Wikipedia facts for their own uses. Wikipedia's authority has become so great that it has become a battleground in which different interests vie for control over the facts that structure our lives.

Understanding how these decisions are made and who Wikipedia benefits most is important for two key reasons. The first is that search engines are recognising Wikipedia as public infrastructure for facts, thus Wikipedia's representations and those advancing such representations need to be critiqued. Facts produced within Wikipedia are increasingly fading into the background and becoming black boxed, so that readers often accept Wikipedia's representation of the world as natural and obvious, when it is actually limited and constructed. Although some well-resourced media organisations and academics are starting to critique Wikipedia's representation of subjects, there is a growing divide between those who understand Wikipedia's representations in the context of its complex rules and affordances, and those who do not. This thesis aims to add to the practical stock of knowledge about how Wikipedia works and for whom it works.

In addition to the identification of Wikipedia as public infrastructure, Wikipedia's *model* is recognised by many as a model for the development of public infrastructure. As more and more private interests are being represented on Wikipedia, it becomes important to understand whether Wikipedia can, indeed, represent public interests as well as

whether Wikipedia's biases are able to be discussed and deemed acceptable or not, or whether they will remain hidden in the black box.

This thesis is a step towards understanding how expert identities are being reconfigured in the current information environment, and what the political implications of such changes are on the representation of knowledges around the world. Initially portrayed as David to traditional media's Goliath, Wikipedia wasn't always recognised as authoritative. Since anyone could edit Wikipedia, it seemed to offer the opportunity for everyone to have their views represented, free from the biases and strains of capitalist media. Wikipedia was part of a growing number of participatory platforms that were considered open for ordinary people to be able to control.

Today it has become increasingly clear that Wikipedia is possibly a new Goliath, reconfiguring the network of actors involved in the production of knowledge about the world. Wikipedia has become a powerful authority over the domain of facts, but questions remain regarding how decisions are made about what to include and exclude in Wikipedia's representation of phenomena, and who (or what) is in charge of such representations. This thesis challenges theories that position Wikipedia and other participatory media projects as foregoing the need for information gatekeepers, giving rise to a group of unnamed amateurs and giving voice to all who participate in their on-going construction. Recognising Wikipedia as a socio-technical phenomenon that has established new centres of authority and expertise through the enactment of symbolic performances, this thesis demonstrates how digital networks give rise to new centres of expertise while reaffirming traditional centres of authority.

Chapter 2: Theoretical framework

(P)eople enact themselves as subjects of power through the Internet, and at the same time bring cyberspace into being (Isin and Ruppert, 2015, p. 12)

Wikipedia has, in the past, been theorised as an example of 'social media' (Fuchs, 2013; van Dijck, 2013), 'peer production' (Benkler, 2006) and 'produsage' (Bruns, 2008). Benkler and Fuchs, for example, classify Wikipedia in terms of its non-profit collaborative model. Benkler (2006) groups Wikipedia with other volunteer-driven projects such as Linux[11] and SETI@home[12] in order to argue that these initiatives represent a new democratic force in society. Similarly, Fuchs (2013) argues that Wikipedia has socialist potential because, unlike exploitative sites such as Facebook and Google, Wikipedia is free of adverts, open source, communally constructed and freely available to all. Van Dijk (2013), on the other hand, considers Wikipedia along with Facebook, Twitter, Flickr and YouTube as a social media platform in which particular cultural, economic and ideological forces sustain an ecosystem of connective media on the Internet.

There is value, however, in theorising Wikipedia in its own terms. Wikipedia defines itself in policy and practice primary as an encyclopaedia, as distinct from sites like

---

[11] Linux is computer operating system assembled under the model of free and open-source software development and distribution.

[12] SETI@home is a scientific experiment that uses computers connected to the internet in the Search for Extraterrestrial Intelligence (SETI). Internet users are invited to participate by running a program that downloads and analyzes radio telescope data (see http://setiathome.ssl.berkeley.edu/).

Facebook or Twitter[13], and specifically as a platform for the representation of facts. Science and Technology Studies, particularly scholarship relating to the communication of scientific facts is therefore relevant to theorising how Wikipedia influences the travel of facts within its socio-technical system and what this means for the distribution of power to represent knowledge.

This chapter provides the theoretical framework for situating the two research questions that have been posed – the first about the reasons why some facts travel further than others in Wikipedia; the second about the impact of such conditions on the ways in which knowledge authorities and experts are being reconfigured. The chapter begins with an outline of the debates within STS about the social vs. technological shaping of technology in order to situate the current research.

The chapter moves on to introducing Morgan's (2010b) framework for factors influencing the travel of facts that includes facts' 'good companions', 'boundaries and terrain' and 'character'. The second part of the chapter highlights the ways in which the boundaries between authoritative and non-authoritative companions, verifiable and unverifiable truth, and reproducible or un-reproducible facts has been defined by knowledge communities in the construction of everyday discourse in the past. Such discourse proves to be both descriptive (whereby rivals denigrate one another using descriptive language, for example) but is also performative in the sense of Isin and Ruppert's (2015) digital speech acts.

---

[13] The policy, WikipediaNotFacebook (Wikipedia:What Wikipedia is Not) declares that 'Wikipedia is not a social networking service like Facebook and Twitter'.

The final section provides an analysis of co-production (Jasanoff, 2004) as a way of framing the second research question regarding authority and expertise. As Jasanoff writes, reconfirming traditional authorities during times of significant socio-technical change is a way of putting things back together, although the identities of the expert will change. Empirically examining how such authorities and identities are being constructed within Wikipedia is highlighted as the core goal of the chapters that follow.

2.1 Social vs. technological determinism

The extent to which technologies and facts develop according to either social or technological forces constitutes a key debate in the field of science and technology studies. On the social constructivism end of the scale, distinctions are made between 'mild' and radical' social constructivism (Sismondo, 1993) with those who represent mild social constructivist views arguing that science and technology involve social factors and radical social constructivism that technologies are purely the product of socially negotiated meanings. The Social Construction of Technology (SCOT) approach developed by Bijker, Hughes and Pinch (1987), for example, recognizes the ways in which social and forces influence science and technology.

> (B)oth science and technology are socially constructed cultures and bring to bear whatever cultural resources are appropriate for the purposes at hand. (Bijker, Hughes & Pinch, 1998, p. 20)

According to the SCOT approach, claims become accepted as scientific truth in three key steps. When scientists initially make claims, and technologies are developed, there is a high degree of interpretive flexibility. Scientific facts can be widely interpreted and technologies variously shaped at this stage of their development. Relevant social groups who are associated with the fact or artifact have different meanings for it because it fulfils particular functions, and solves specific problems for them. These meanings

inevitably come into conflict with one another as they suggest different alternatives for the fact or the artifact.

A single truth emerges or technologies are stabilized when consensus emerges and controversies are terminated. This process of closure and stabilization occurs through either rhetorical closure (whether relevant social groups see the problem as being solved) or according to a redefinition of the problem within a wider context.
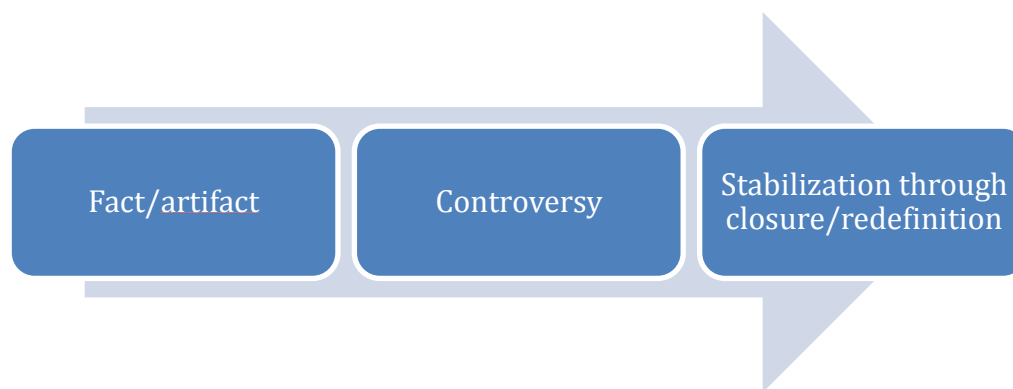
Figure 2.1 The stabilization of facts and artifacts according to Bijker, Hughes & Pinch (1987)
Source: Adapted from Bijker, Hughes & Pinch (1987)

MacKenzie & Wajcman (1985) reiterated the importance of the social forces underpinning the ways in which science and technologies develop in their influential volume titled 'The Social Shaping of Technology' (MacKenzie & Wajcman, 1985). The authors stressed the need for recognizing how technology doesn't affect society in some independent way but that technology is also socially shaped. According to MacKenzie and Wajcman, it is not that technology has no social effects but that the effects of technology are complex and contingent (MacKenzie & Wajcman, 1985, p. 4). Included in the volume were articles such as Langdon Winner's 'Artifacts have politics' that indicated how technologies are not neutral but are designed in order to open certain social options and close others.

In addition to recognising the importance of social and cultural factors in influencing the development of technology, MacKenzie and Wajcman noted in the second volume (1999, p. 23) that they had originally neglected the influence of technology upon social relations and argued that it is important to recognise that social relations are not independent of technology. They explain that actor-network theory, developed by scholars including Latour, Callon, Akrich and Law is useful in understanding how technology and society are mutually constituted, that artifacts and technologies are what makes society possible and that both society and technology are made of the same 'stuff': networks linking humans and non-humans together.

Actor network theory (ANT) proposes that social issues or practices must be recognized as networks of relationships among human actors and non-human actants. ANT scholars reject the classification of issues or practices as either nature or culture, science or politics, but rather suggest that material artifacts can be agents within networks of relations. Although the concept that artifacts can have agency is controversial, the heterogeneity of actor networks places artifacts and objects centrally within the theory.

According to MacKenzie (1996), what is important about ANT is that it reminds us to keep two aspects of technical change in mind: the first is the way in which the physical aspects of heterogeneous engineering are influenced by the demands of its social aspects, that is the 'social shaping of technology'; the second is that 'artifacts have politics' (Winner, 1980), that is, technologies are not neutral; their adoption and operation often involves changes to that order.

> (T)he actor-network perspective offers a useful critique of the fact that much social theory conceives of social relations as if they were simply unmediated relationships between naked human beings, rather than being made possible

and stable by artifacts and technologies. (MacKenzie, 1996, p. 14)

Actor Network Theory proposes that what we define as social issues or phenomena are actually networks of human actors and non-human actants interacting together (and often at odds with one another). Michel Callon (1986), one of the proponents of ANT, highlights the ways in which controversies are resolved through the example of the controversy about the causes for the decline in the population of scallops in St. Brieuc Bay in France. Callon describes how the researchers who participated in the debate about what were the causes of the decline in the population of scallops each developed contradictory arguments and points of view which lead them to propose different versions of the social and natural worlds. He demonstrates how the researchers who were successful in imposing their definition of the situation on others were those who were able to define their identities in such a way that they became an obligatory passage point in the network of relationships that they were building. Furthermore, successful researchers were those whose determinations of the roles and identities of different actors were enacted and unopposed. The fishermen of St. Brieuc Bay, the researchers' scientific colleagues, the scallops of St. Brieuc Bay and the researchers themselves become enrolled in a representation of the controversy according to the researchers' own interpretation of actors' interests and behavior.

A process of translation occurred when the researchers imposed themselves and the definition of the situation onto others through problematisation, interessement, enrolment and mobilization. Problematisation refers to the movement in which certain actors become indispensable in the network; interessement, the processes by which actors lock other actors into the roles that they propose; enrolment as the device by which a set of interrelated roles is defined and attributed to actors who accept them,

52

and mobilization, a set of methods used by actors to ensure that supposed spokespeople for various relevant collectivities are properly able to represent those collectivities.

> Translation is the mechanism by which the social and natural worlds progressively take form. The result is a situation in which certain entities control others. Understanding what sociologists generally call power relationships means describing the way in which actors are defined, associated and simultaneously obliged to remain faithful to their alliances. (Callon, 1986, p. 224)

Latour and Woolgar's 'Laboratory Life' (1979) similarly analysed the ways in which science was being constructed within another actor network, that of the laboratory. In particular, Latour and Woolgar recognized the role of the material environment of the laboratory in influencing the kinds of facts that are created. Investigating patterns of communication in a scientific laboratory, the authors traced the routine practices of publishing papers, negotiating research finances and seeking scientific prestige in the trajectory of scientific facts. Latour and Woolgar note that there is a heavy focus on documents and literature in the laboratory with scientists spending 'the greatest part of their day coding, marking, altering, correcting, reading, and writing.' (p. 49)

Latour extended these findings into his book, 'We have never been modern' (1993) where he proposes that in order to study the relationship between science and society, we need to maintain the unity between science and the social. Modern (Western) society has constructed this division but it continuously threatens to unravel with the existence of what Latour calls 'hybrids' that mix politics, technology, science and culture. Objects such as the cloned sheep Dolly or government experts are neither natural nor social but rather linked together in heterogeneous actor-networks constituted by the public interaction between people, heterogeneous objects and discourses.

According to Sheila Jasanoff (2004), however, the problem with Latour's conception of power within actor-networks is that it is determined by the size and concentration of

the network. Latour writes that power tends to concentrate in 'centres of calculation' such as the printing presses, statistical formulas, maps and other 'inscription devices' which render dominant perceptions of the world into portable representations (Latour, 1987, 1992, 1993). There are significant gaps in Latour's articulation of power relations according to Jasanoff because he does not account for diversity in the acceptance of scientific claims by some groups as opposed to others.

> [Latour is] silent on why technological practices or the credibility of scientific claims varies across cultures; why some actor-networks remain contested and unstable for long periods while others settle quickly; why work at some nodes stabilizes a network more effectively than at others; or what role memories, beliefs, values and ideologies play in sustaining some representations of nature and the social world at the expense of others. (Jasanoff, 2004, p. 39)

Instead, Jasanoff argues for a co-productionist approach to power in which power becomes lodged in representations, discourses, identities and institutions during times of significant socio-technical change. Underlying the theory of co-production is a theory of the impact of social and technological change that sees both social and technological factors as imbricated with one another, rather than happening in sequence. Instead of technologies causing social change, technologies are themselves created as a result of changes that cannot be narrowed down as either social or technical.

Co-production can be seen as an attempt to bring together strains of social and technological determinism over the history of STS, as well as to bring in a normative lens that recognizes the role of power in socio-technical change which has been lacking in the field.

Jasanoff's expansion of the co-production idiom focuses on a Foucauldian notion of power where power is not about the ability to conduct violence or force, but rather in the sense in which power is the ability to classify. Foucault (1980) argues that society

54

tends to value particular forms of knowledge over others, to recognise some forms of knowing and not others, to accord some people with the designation of expert and others as lacking true knowledge. The power to represent knowledge is therefore the power to define the rules underlying determinations of what is true and false. Foucault uses the term 'power/knowledge' to signify how truth is constructed by societies according to the power of certain groups.

> Truth is a thing of this world: it is produced only by virtue of multiple forms of constraint. And it induces regular effects of power. Each society has its regime of truth, its "general politics" of truth: that is, the types of discourse which it accepts and makes function as true; the mechanisms and instances which enable one to distinguish true and false statements, the means by which each is sanctioned; the techniques and procedures accorded value in the acquisition of truth; the status of those who are charged with saying what counts as true. (Foucault, 1980, p. 131)

Foucault explains how these 'regimes of truth' are the result of scientific discourse and institutions and are reinforced and redefined continuously through the media, the education system, political and economic ideologies. The battles that are waged over truth are not only about which are the absolute truths but rather about the 'rules according to which the true and false are separated and specific effects of power are attached to the true'; a battle about 'the status of truth and the economic and political role it plays' (Foucault, 1980, p. 132). 'Truth regimes', in others words, are sets of rules that define truth in particular ways.

Scientific objectivity is a 'truth regime' that Donna Haraway (1988) critiques in her work on 'situated knowledges'. According to Haraway, objectivity as it is practiced in male-dominated science is an illusion. There can be no 'infinite vision' – it is a 'god trick' (Haraway, 1988, p. 581). On the other hand, recognising the situated nature of all knowledge enables a new understanding of objectivity which is about 'limited location and situated knowledge, not about transcendence and splitting of subject and object'

(Haraway, 1988, p. 583). Haraway believes that such 'subjugated standpoints' are preferred 'because they seem to promise more adequate, sustained, objective, transforming accounts of the world.' (Haraway, 1988, p. 883) From this perspective, all knowledges are local knowledges; there can be no universal knowledge.

The benefit of Jasanoff's co-production idiom, according to Lievrouw (2014, p. 31) is that it is useful in understanding the relationship between technology and society, in particular how power is materialized not as 'an abstract "force" or institutional "structure," but… as observable in the physical forms of social practices, relations, and material objects and artifacts' (Lievrouw, 2014, p. 31). Citing Irwin (2008, p. 589), Liewrouw writes that,

> Co-production encourages a move away from strong social determinism, and the assumption that 'social controversies around science are "really" all about politics or that complex areas of innovation can be reduced to "social" construction' (Lievrouw, 2014, p. 31)

According to the co-production idiom, the making of identities is one of the key ways in which technology and society co-produce one another. Expert identities are the result not only of the mastery over particular skills, but the tacit knowledge and authority accorded to particular identities by society. As such, it is important to understand how scientific authority involves the power to define what counts as the truth and how scientific authority is conferred upon by society.

Underlying the theory of co-production is a theory of the impact of social and technological change that sees both social and technological factors as imbricated with one another, rather than happening in sequence. Instead of technologies causing social change, technologies are themselves created as a result of changes that cannot be narrowed down as either social or technical.

Co-production can be seen as an attempt to bring together strains of social and technological determinism over the history of STS, as well as to bring in a normative lens that recognizes the role of power in socio-technical change which has been lacking in the field. Although co-production is useful in framing identity as a co-product of society and technology, it has mostly been applied to cases in which there are already stable categories of science and scientist, and there is consequently a gap in understanding the particularities of contexts in which activity is almost entirely mediated by software and code. Mediated environments afford very different types of activity, identity and power relations.

Users are operating under varying levels of anonymity and pseudonymity that provide different levels of identity construction and certification; users have different levels of technical permissions that afford different types of activities; and the architecture of the spaces in which users operate are determined by a particular type of materiality: the materiality of code. These materialities have an effect on the kinds of identities that are either re-affirmed or co-produced because they affect which actors are able to undertake which actions in particular environments.

In addition, the process of fact building in particular is necessary to be explored in order to understand both the social and technical forces that are relevant to the travel of facts within Wikipedia's socio-technical structure. In the next section I highlight research relating to the travel of facts, followed by analyses of how the environment in which facts travel is negotiated and architected.

2.2   Facts and knowledge

Although Wikipedia aims to represent 'the sum of all human knowledge' (Wikimedia Foundation, n.d.), one of the site's foundational policy documents (Wikipedia:What Wikipedia is Not) indicates that Wikipedia is an *encyclopaedia* that represents *facts*, rather than a series of other types of formats, media or systems. There are three key differences between knowledge and facts that provide insight into some of the foundational limitations that Wikipedia encounters when trying to represent knowledge about the world.

Firstly, facts are a representation of knowledge rather than a mirror of reality. Reality must be represented and communicated in language and, according to social psychologist Sandra Jovchelovitch who writes about knowledge in context, representation is the only path to knowledge that we have.

Our knowledge of the world depends on representational processes; as mediating structures bridging the world of subjects and the world of objects they deeply affect the structure of knowledge. Knowledge in this sense is neither a copy of the world nor is it the world but it is in the world. Knowledge systems are proposals of the world – literally representations – whose processes of construction we need to understand and to unpack if we are to understand their complexity and variability in social life (Jovchelovitch, 2006, p. 100).

Secondly, the social features of facts are progressively removed as they travel to other contexts. Whereas knowledge resides in the mind or in the flow of practice (Lave & Wenger, 1991), facts are structured and perceived as distinct from the context from which they arise and can therefore be moved around from one context to another.

58

Jovchelovitz writes that, although 'knowledge is a plural and plastic phenomenon', for many 'the trick behind the construction of true knowledge seems to be the progressive detachment of the internal structures of knowing from the subjects, communities and cultures that give knowledge its substance and its raison d'être' (p. 98).

Thirdly, there are different ways of knowing the world, but only certain groups and identities are attributed with being able to develop true knowledge in the form of facts, whereas other groups are said to merely produce opinion, myth, or belief. Jovchelovitz argues that it is in the realm of science that contemporary societies tend to locate true knowledge in the form of facts and that people other than scientists ('a housewife, a five-year-old child, or a peasant living in a rural community') are determined to rather hold 'lay beliefs, ideologies, myths, or superstitions, but not knowledge' (p. 99).

In order to understand statements that are taken as factual and therefore authoritative, we need to understand the process by which facts travel from their point of origin to other contexts. The travel of facts is critical to their power; facts that don't travel well cannot become known and therefore do not have power to define their phenomena. Facts need to move out of the laboratory or other origins of their production in order to become known, independent and authoritative. According to Steven Shapin, the communication (or representation) of facts is not distinct from, but part of their construction.

> (S)peech about natural reality is a means of generating knowledge about reality, of securing assent to that knowledge, and of bounding domains of certain knowledge from areas of less certain standing. (Shapin, 1984, p. 481)

In the context of Wikipedia, where editors are forbidden from producing their own research, all facts represented on Wikipedia are produced elsewhere. Like any other

representation of knowledge, representations of facts on Wikipedia are not a perfect mirror but rather a (re)presentation of knowledge with its own unique characteristics. In communicating facts through the medium of the Wikipedia article and in response to Wikipedia's policies, norms and local interactions with other editors, Wikipedians add particular elements to the fact as it moves through the Wikipedia system.

The way in which Wikipedians summarise a fact in a scientific article, for example, imposes a particular emphasis and classification, a particular association as the fact is grouped with other facts. The fact needs to remain largely intact as it travels, but its companions and the environment in which it travels is instrumental to that travel and therefore to its on-going construction.

A research project entitled 'How Well Do "Facts" Travel?' (Morgan, 2010a) investigated the transmission and reception of facts 'traveling across time, between disciplines, between academia and policy, between the lay public and the specialist professional, and in the physical sense across countries, embodied in people, spoken, written, performed and executed, or disembodied in books, diagrams, and technologies' (see the project website at http://www.lse.ac.uk/economicHistory/Research/facts). In an introduction to the book from the project, Mary S. Morgan (2010b) defines facts as 'autonomous, short, specific and reliable pieces of knowledge' (p. 8). Facts come in many forms. They are distinguished by their diversity in expression, scope and size by the different communities that employ them. Facts can be expressed and represented both linguistically or in the form of objects and artifacts, they can be little, big, singular, multiple or generic (p. 8).

The travel of facts is the central remit of the research project, with travelling 'well' being defined in two key ways. Firstly, when a fact travels well it travels 'with integrity' so that it travels more or less intact from one place to another. Secondly, a fact travels well when it travels 'fruitfully', that is, beyond its spatial or disciplinary boundaries to find new users and uses of the fact (p. 12). On the other hand, facts may be impeded from traveling well by 'bad companions, companions who alter the fact to subvert it, re-label it, cast doubt on it, and otherwise discredit it as they see it on its way' (p. 30).

Expanding on the 'fruitful' travel of facts, Morgan articulates their position with regards to the construction or discovery of facts that has been a feature of debates in the sociology of knowledge and the so-called 'science wars' of the late twentieth century. According to Morgan, there are two core ways of thinking about facts from the production side. On the one hand, there is the idea that facts are found or discovered after heavy labour in the laboratory, field, archive or museum. On the other hand, facts are viewed as being constructed out of activities of social networks and practical instruments. Since Morgan is focused specifically on the travel of facts after production, she looks to two theoretical positions regarding whether (or how) facts change as they travel.

According to Latour (1987), the 'marks' of science are mobile and will travel only if they are immutable, presentable, readable and combinable. Fleck (1979), on the other hand, argues that facts are developed and understood only within knowledge communities. As knowledge travels from one community to another it has to be translated and in the process changes to some degree its meaning and thus loses its integrity in traveling.

Morgan argues that her sense of what happens to facts as they travel 'fits untidily between Latour's marks and Fleck's community facts' (p. 13). Their notion of facts is that they are not always as mobile as Latour would suggest since they are not necessarily linguistic (marks on a headstone, for example). Their conception is that facts are also more independent than Fleck suggests facts are. According to Morgan, the travel of facts may involve the transformation in meaning since a community defines its own facts, but the circulation of such knowledge doesn't *necessarily* involve its transformation (p. 15).

Morgan uses the analogy of facts as rubber balls to articulate the changeability of facts as they travel:

> They have a certain shape; they can be carried, rolled, squeezed, bounced, kicked and thrown without harm to them; and they can be used in many different ways and in different situations. (Morgan, 2010b, p. 15-16)

In addition to their steadfast and mutable characteristics, facts have accompanying details or qualifications that constitute contextual elements that may be shed on a fact's travels. In addition to losing contextual elements, facts may also pick up extra elements on their travels and become sharpened (p. 16).

In the context of a Wikipedia article, facts are reflected in the article's title and its redirects[14], in the content of the article and the associations of facts with particular headings, in categories and images and captions and infoboxes and hyperlinks. A fact travels well within Wikipedia if the source is accurately reflected in the context of the article, if the fact is translated into other languages on Wikipedia and linked to other

---

[14] A redirect is a page that has no content itself but sends the reader to another page, usually an article or section of an article. Redirects enable a page to be reached under alternative titles on Wikipedia.

articles. The fact travels well if it and its constituent facts were extracted so that they could be represented in Wikidata, and other platforms that rely on Wikipedia data. A fact travels less well within Wikipedia if the source is inaccurately summarised in the article, if it is not translated into other languages, if its viability as a reliable fact is questioned by other editors using disruption mechanisms and if the facts were not extracted by Wikidata and thus not available for extraction in other platforms such as Google.

Facts are present in every domain, and in every domain, facts are contested. Although much research in STS has focused on the construction and flow of facts in the field of science, facts are the bedrock of both academic and non-academic enterprises that are based on the creation and dissemination of facts and knowledge. Like natural science, history advances by scholars interrogating facts established in the past. Facts are a form of shared knowledge. Facts do not represent belief or opinion but rather knowledge that,

> according to (a community's) standards of evidence of discipline, time and place, that community has good reason to take those things as facts and will be likely to have the confidence to act upon them as facts. (Morgan, 2010, p. 11)

Facts are produced according to these standards of evidence imposed by the community such as rules of observation, experimentation or inscription in scientific discipline. In order for facts to become accepted, however, they have to communicated to others.

According to Morgan, it's not that there is a free market for facts and that 'good facts' will necessarily crowd out 'bad facts'. Actually, some facts find fierce resistance among people with opposing interests and ideologies. Morgan highlights three key factors that impact the travel of facts. They include 'good company', 'boundaries and terrain', and 'character' (see figure 2.2).

Traveling companions	Boundaries and terrain	Character
Authorities	Norms	Attributes
Labels and packaging	Policies	Characteristics
Technical standards	Epistemologies	Functions
Processes		

Figure 2.2 Factors influencing how far a fact travels according to Morgan (2010b)

Source: adapted from Morgan (2010b)

'Good company' refers to the people and structures that support a fact's travels but are not part of the facts themselves and can be discarded when the fact reaches a new destination. According to Morgan, good travelling companions are required for facts to travel well and with authority.

> (Travelling companions) range from the mundane level of labels and packaging, to the more material vehicles of transportation, as well as to the people involved in chaperoning, and from the various kinds of institutional structures that support travelling knowledge, to the technical standards that carry facts with them (Morgan, 2010b, p. 27)

The packaging of facts ensures that they travel well in terms of retaining their integrity as they pass from one location to another. Facts traveling in images, in specimens, in bio-informatics data all need to be carried – by people, by data processes and technical standards. This is not about the agencies of individual producers of facts but those who package facts for travel, the users or audiences who unpack them, the network of people and things via which they travel and the social arrangements within which these travels are embedded (p. 21)

In Wikipedia's socio-technical environment, 'traveling companions' can be identified in the editors who summarise sources and reflect them as facts in articles, by the classification standards and techniques operating within Wikipedia and its associated projects including Wikidata, as well as the technical permissions that enable certain editors greater control over editing all aspects of the article, including the tools used to frame content such as templates.

'Boundaries and terrain' refers to the disciplinary landscape, material elements, and requirements for a specific technical understanding that limit the range of traveling facts. According to Morgan, the terrain metaphor can be defined in terms of numerous features.

> We can construct the terrain in *sociological* terms, for example as a disciplinary landscape in which expertise, trust, and power form the features of the terrain and define the barriers to be overcome. Or we can construct it in terms of the *material elements* of the science or humanities in which models, instruments and experiments - or archives and previous historical authorities - constitute the terrain. A third possibility is to interpret the terrain and boundaries in *cognitive* and *epistemic* terms, where the requirements for a specific technical understanding, or a knowledge of historical period, limit the range of the travelling facts or their ability to remain intact as they travel. (Morgan, 2010b, p. 31)

Applied to Wikipedia, I define boundaries and terrain as the norms, policies and epistemological standards applied within Wikipedia's environment that enable certain facts to be travel well and others to be stopped in their tracks. How Wikipedia defines its boundaries is a product of its encyclopaedic identity, its origins in the free and open sources software movement and the Internet as liberatory, and its relationship to authoritative sources of knowledge. This environment plays a very particular role not only in how facts travel through it, but which facts will travel more easily than others based on how they are defined by different actors as being located within or outside of the boundaries being set.

Finally, the character of facts refers to the specific attributes, characteristics, and function of facts. Facts have unique qualities with the potential to develop their scope or become generic. Facts can be 'understandable', 'surprising', 'colourful', 'reproducible', and 'adaptable'. Facts with character travel well, that is, they travel far and with integrity.

Within Wikipedia's socio-technical system, certain facts travel further than others because of their character. Thus, facts that are reproducible such as demographic data about cities are well represented in Wikipedia. Reproducibility depends on the availability of data about cities as well as the automated agents on Wikipedia that produce and translate particular types of articles and facts. Not all countries in the world maintain the same level of detail for cities and municipalities, towns and villages, and so not all cities' are represented by facts that are reproducible.

Morgan produces a clear framework for understanding how facts travel, but research into the construction and popularization of facts has shown that the process of defining who are 'good' traveling companions, what are the boundaries between fact and opinion, as well as which facts are more reproducible than others are equally important to understanding the travel of facts. Studies in this arena have considered the social (Bijker, Pinch & Hughes, 1987) and socio-technical (Latour, 1987; Jasanoff, 2006; Isin and Ruppert, 2015) interactions that result in certain facts being spread further than others, the ideological impetus behind the travel of facts (Geertz, 1973; Gieryn, 1983, 1999) and the infrastructure required for the spread of facts (Shapin, 1984).

Common to all of these studies is a conception of rhetoric as the basis of definitional practice and boundary work. Studies have predominantly been done in the realm of science studies where scholars have investigated how stakeholders collectively articulate the boundaries between facts and opinions, science and non-science, and ideology and science. Also relevant are studies of controversies where particular facts are being debated through debates over who are the appropriate authorities to speak on behalf of particular subjects.

2.3 Language, ideology and rhetoric

Clifford Geertz (1973) developed a theory of ideology out of a concern that ideology was becoming a pejorative term rather than an analytical concept. According to many scholars, ideology occurs when there are deviations from social scientific objectivity and ideology is frequently used as a pejorative term for statements that the speaker believes to be biased, emotional or distorted as opposed to independent, unaffected and scientific. Geertz argues that the boundary between ideology and science is unclear.

> Where, if anywhere, ideology leaves off and science begins has been the Sphinx's Riddle of much of modern sociological thought and the rustless weapon of its enemies. (Geertz, 1973, p. 194)

According to Geertz, what is missing is a way of understanding how some ideological statements achieve resonance and success, while others are determined to be lies or propaganda. Instead of evaluating statements according to whether they are ideological or scientific, we need to separate out the concept of ideology and understand why certain statements are amplified while others are stopped in their tracks.

Two key theories dominate our understanding of ideology, according to Geertz. The first is the 'interest theory'; the second, the 'strain theory'. Interest theories of ideology, advanced by Marxist scholars, argue that ideological statements are expressions of

individuals' strategic moves to garner advantage in the social (class) system. Strain theory, on the other hand, frames ideology as a response to the inherent inconsistencies (strains) of everyday life.

> For the first, ideology is a mask and a weapon; for the second, a symptom and a remedy. In the interest theory, ideological pronouncements are seen against the background of a universal struggle for advantage; in the strain theory, against the background of a chronic effort to correct sociopsychological disequilibrium. In the one, men pursue power; in the other, they flee anxiety. (Geertz, 1973, p. 201)

Interest and strain theories are not incompatible to Geertz. Whereas interest theories represent strains on the individual, strain theories represent strains on society. Both are symbolic, both see ideology as distorting social reality, both require that in order to explain ideological statements on needs to examine their social context.

Geertz provides a definition of ideology as 'systems of interacting symbols, as patterns of interworking meanings' (p. 207). People become attached to certain symbolic systems rather than others, argues Geertz, and analysts require a deeper understanding of how symbolic formulation works in order to understand their meaning. We need, according to Geertz, an understanding of style – something that is currently missing from the sociologist's toolbox.

> With no notion of how metaphor, analogy, irony, ambiguity, pun, paradox, hyperbole, rhythm, and all the other elements of what we lamely call "style" operate – even, in a majority of cases, with no recognition that these devices are of any importance in casting personal attitudes into public form, sociologists lack the symbolic resources of out of which to construct a more incisive formulation. (Geertz, 1973, p. 209)

In order to understand meaning making, Geertz recommends that we look at the social realities in which the symbol appears in order to understand how socio-psychological strains are expressed in symbolic forms. Symbolic systems are frameworks for understanding how to act in the world; they are 'extrapersonal mechanisms for the

perception, understanding, judgment, and manipulation of the world' (p. 216). Geertz writes that during times of significant change we have the most use for ideologies in order to make our way.

> It is precisely at the point at which a political system begins to free itself from the immediate governance of received tradition, from the direct and detailed guidance of religious or philosophical canons on the one hand and from the unreflective precepts of conventional moralism on the other, that formal ideologies tend first to emerge and take hold. (Geertz, 1973, p. 219)

According to Geertz, this doesn't mean societies return to naïve traditionalism when faced with significant change but rather 'ideological retraditionalization' (p. 220). It is when socio-psychological strain meets with an absence of cultural resources by which to make sense of the strain that ideologies are formulated and advanced.

Gieryn (1983, 1999) adopts Geertz's definition of ideologies as systems of interacting symbols to indicate the main ways in which the boundaries between science and non-science are determined. Studying the rhetoric engaged in by scientists in ousting their rivals during public debates, Gieryn writes that boundary-work is used by ideologists of a profession or occupation such that:

> when the goal is expansion of authority or expertise into domains claimed by other professions or occupations, boundary-work heightens the contrast between rivals in ways flattering to the ideologists' side; when the goal is monopolization of professional authority and resources, boundary-work excludes rivals from within by defining them as outsiders with labels such as 'pseudo',' 'deviant', or 'amateur'; when the goal is protection of autonomy over professional activities, boundary-work exempts members from responsibility for consequences of the work by putting the blame on scapegoats from outside. (Gieryn, 1983, p. 792)

Scientists' authority, according to Gieryn, is maintained by the public articulation of scientific practice and episodes of boundary-work which involves attributing particular characteristics to the institution of science in order to construct 'a social boundary that distinguishes some intellectual activity as non-science' (p. 782). Gieryn theorised that scientists establish boundaries between science and non-science, scientists and non-

69

scientists through a process of classification work in which they tried to show how science was unique from rival forms.

Gieryn acknowledges that the characteristics of science vary - science is no single thing. 'The boundaries of science are ambiguous, flexible, historically changing, contextually variable, internally inconsistent, and sometimes disputed' (p. 792). The characteristics assigned to science are sometimes inconsistent because scientists need to erect separate boundaries in response to different obstacles to their pursuit of authority and resources. Boundaries are sometimes contested by scientists with different professional ambitions and there is ambiguity from the simultaneous pursuit of separate professional goals, each requiring a boundary to be built in different ways. Whereas some speakers use certain characteristics to define science according to their own interests and strains, others may attribute science with completely different characteristics.

Furthermore, success in boundary-work, according to Gieryn, is temporary. Success, defined as the successful attribution of particular features of science to the field, is always provisional and contextual. Instead of operating as 'determinants of who wins', 'Scientific practices and antecedent representations of it form a repertoire of characteristics available for selective attribution on later occasions' (Gieryn, 2001, p. 406). Descriptions of science and scientists, in others words, operate as specific, concrete practices that can be more easily pointed to in order to advance interests. Such descriptions act as rhetorical evidence or maps of knowledge work.

Gieryn's understanding of how struggles to define science are won and lost is useful because it points to the ways in which success is determined by the ability to define particular roles using rhetorical moves that set up a territorial arrangement and have those definitions accepted. Strategies, then, are particular to the context and the

environment; they require situationally specific vocabularies in order to succeed. Although Geertz and Gieryn predominantly talk about boundaries between science and other knowledges, the theory of ideologies as symbolic systems and boundaries being determined by ideological statements or rhetoric can be applied to other knowledge communities.

The problem with Gieryn and Geertz's conception of the methods used to construct boundaries between different types of knowledge and different types of knowers is that they focus only on the descriptive nature of language instead of also recognizing that language can also be performative (Austin, 1975). Knowledge claims aren't just made, they are made to perform; the utterance of statements does work in the world.

Shapin (1984) recognises the importance of literary and social conventions, in addition to material technologies required to legitimate knowledge. Shapin recounts what was required in order for Robert Boyle's experiments in pneumatics in the late 1650s and early 1660s to succeed in being seen as the legitimate means by which knowledge was to be generated and evaluated. Infrastructure needed to be constructed in order to enable the authentication of factual claims. Such infrastructure included the material technology of the air-pump that produced the experiments, a literary technology whereby the phenomena produced by the pump were made known to those who weren't direct witnesses, and the social technology of conventions people should employ in dealing with one another and considering knowledge claims. Boyle succeeded in building, exemplifying and defending the technology necessary for producing legitimate knowledge by developing material, social and literary technologies.

Shapin recounts how Boyle suggested that matters of fact were to be produced in a public space where experiments were collectively performed and directly witnessed and an abstract space constituted through virtual witnessing. Virtual witnessing was enabled by the use of literary techniques that produced realistic images of the scene of the experiment with great deal of circumstantial detail and using prolixity and iconography. An authenticated matter of fact was treated as a mirror of nature; a theory, by contrast, was clearly man-made and could be contested. Boyle's linguistic boundaries thereby acted to segregate what could be disputed from what could not. Shapin argues that literary conventions appropriate for virtual witnessing needed to be constructed in order to support consensus-building, reconciliation and scale.

> An appropriate language had to perform several functions. First, it had to be a resource for managing dissent and conflict in such a way as to make it possible for philosophers to express divergent views while leaving the foundations of knowledge intact, and, in fact, buttressing these foundations... Second, it had to facilitate reconciliation amongst existing sects of philosophers, mobilizing that reconciliation so as to reinforce the foundational status of matters of fact... Third, such a language had to constitute a vehicle whereby matters of fact could effectively be generated and validated by a community whose size was, in principle, unlimited. (Shapin, 1984, p. 507)

The product of the pump was not an inscription as Latour and Woolgar (1979) had argued; it was, instead, a visual experience that had to be transformed into an inscription by a witness. Witnesses were chosen among the ranks of professors rather than peasants and followed taken-for-granted conventions for reliability. Membership was also limited to those who could use the appropriate conventions in conducting their 'virtual witnessing'.

> Not everyone may speak; the ability to speak entails the mastering of special linguistic competences; and the use of ordinary speech is taken as a sign of non-membership and non-competence. (Shapin, 1984, p. 509)

Latour (1987) demonstrates the importance of literary conventions of fact building within modern science where scientists use tools such as literature, numbers, images

and figures in order to convince others of their claims. According to Latour, although the field of rhetoric is 'despised' because it mobilises external allies such as 'passion, style, emotions, interests, lawyers' tricks and so on' in the service of an argument, rhetoric is the primary way in which scientists convince others of their claims (p. 61). The difference between rhetoric and science, according to Latour, is not that rhetoric makes use of allies that science refrains from using but that rhetoric 'uses only a few and (science) very many' (p. 61). External allies (or resources) for scientists come in the form not only of style but also in the form of literature, numbers, geometrical figures, equations, images, mathematical objects etc.

Latour's study highlights the ways in which the performance of conventions is part of what it means to be a scientist, a part of the profession's identity. He outlines how scientists must continually work to build the number of nodes in their network by 'bringing friends in' (by alluding to what other authorities said or wrote), referring to former texts (by either supporting or attacking them), and being referred to by later texts since a claim or statement needs to be cited by the next generation of papers in order to become taken for granted (p. 30-44). In order to persuade, scientists initially connect their claims with the claims of other, more established writers and their writings, and then eventually remove those citations as a fact becomes taken-for-granted.

The value of Shapin's focus on literary conventions and Latour's highlighting of the actions taken by scientists in the development of facts is their recognition of the performativity of language. Language does not only act to describe the world; language can also perform. J. L. Austin (1962) developed the concept of 'performative utterances' to describe the types of speech acts enabled by promising, greeting, warning, ordering

and congratulations statements, for example. Since then, the field of linguistics and the philosophy of language have investigated how speech acts are distinguished by the intentions behind the speech act: 'there is the act of saying something, what one does in saying it, such as requesting or promising, and how one is trying to affect one's audience' (Bach, 1998, p. 81).

Isin and Ruppert (2015) extend the theory of speech acts to encompass a theory of digital speech acts. According to the authors, digital speech acts involve conventions that have been developed in online environments that include not only words but also images and 'sounds and various actions such as liking, coding, clicking, downloading, sorting, blocking, and querying' (p. 13). These constitute a type of speech act that has performative (and political) force. In particular, digital speech acts, according to Isin and Ruppert, are acts in which a digital citizen subject articulates that either 'I, we, (or) they have a right to' (p. 12) perform these acts.

The authors base their theory of digital performativity on Foucault's notion of power as 'action upon action' or 'conduct upon conduct' where individuals carve out social positions through the iterative enacting of conventions over time. Some speech acts have performative force: we don't only talk about the world, but we also act through talking, and in the online space, we talk through doing.

Isin and Ruppert argue that 'Digital citizens come into being through the meshing of online and offline lives' (p. 19) and that cyberspace is not separate from 'real' space, but that it is 'a space of relations between and among bodies interacting through the Internet' (p. 12). Despite previous debate about the unproductiveness of the term 'cyberspace' (Graham, 2012), Isin and Ruppert argue that the term is useful in

demonstrating how 'people enact themselves as subjects of power through the Internet, and at the same time bring cyberspace into being' (p. 12).

Similarly, Dodge and Kitchin (2005) argue that space needs to be theorized as ontogenetic rather than as absolute. An absolute theory of space is that it is 'a geometrical system of organization, a kind of absolute grid within which objects are located and events occur' (Dodge & Kitchin, 2005, p. 171) but Dodge and Kitchin argue that this is reductionist and depicts space as 'natural and given' (Dodge & Kitchin, 2005, p. 171). Instead, they argue for an ontogenetic reading of space where space is understood 'as continually being brought into existence through transductive practices (practices that change the conditions under which space is (re)made)' (Dodge & Kitchin, 2005, p. 162). In this sense, space is given meaning through human endeavor and produced through social relations. The authors argue that the

> objects, infrastructures, processes and assemblages (produced by code)… engender, transduce space – beckon(ing) new spatial formations and spatiality into existence. (Dodge & Kitchin, 2005, p. 171)

Liking a post on Facebook, for example, constitutes a digital act mediated through the network in which real bodies need to be present. In Wikipedia's terms, such actions involve the vast set of conventions that have been developed by users over time including: warning, proposing for deletion, amongst others. These acts constitute a means of social struggle among actors with a variety of goals, interests and intentions.

Digital acts signify four political questions about the Internet according to Isin and Ruppert (2014). They include anonymity, extensity, traceability, and velocity (p.13). The Internet's affordances or materiality are an important feature that affects socio-technical relations and the representations, discourses and identities that have been

produced as a result of such mediation. The study of software demonstrates that, not only are digital speech acts constitutive of a new type of language known to those who act, but digital speech acts can also be multiplied because of software's capacity for secondary agency through automated processing.

2.4 Software and new tools for language construction

Coded environments are virtual in the sense that they are constructed out of bits and bytes rather than bricks and mortar, but digital environments have stability and materiality despite the fact that they exist in a virtual form. Although digital, code has effects in the world; code enables certain behaviour while preventing other behaviour. Code is the stuff out of which identities are created, the spaces in which knowledge and expertise are enacted and the tools that are wielded by different groups in the construction of ideology.

The materiality of the coded environment is important to the construction of ideologies necessary to maintain the boundaries between expert identities. Although the concept of materiality has been widely defined in multiple disciplines, I adopt Lievrouw's (2014) definition in relation to communication and technology studies as 'the physical character and existence of objects and artifacts that makes them useful and usable for certain purposes under particular conditions' (p. 25). The focus on artifacts enables a consideration of how the objects of digital environments are critical to understanding practice, but this requires an understanding of how digital objects have materiality.

Faulkner and Runde (2013) recognise that the defining features of objects that they endure and that they are structured. Objects differ from events in that they are the same at each point in time in which they exist and they are structured in that they are

composed of distinct parts. Each object can be composed of other objects and an object's constituent parts aren't always the same over time, for example when living organisms change over time. The authors go on to show how technological objects are identified according to their function or use for a particular group of people.

The difference between physical objects (for example, a bicycle) and digital objects (such as an algorithm) is that digital objects don't possess spatial characteristics (such as location, shape or volume) but instead consist of symbols that are organized to conform to the rules of the language in which they are expressed (Faulkner & Runde, 2013). Technological identity suggests that identity flows from both function and form. In other words, the function of a technological object flows from that which is ascribed to an object by particular social groups, and in order for a function to be sustained over time, an object must possess physical capabilities. Human activities and social structure are recursively organised so that we do things either consciously or through routine, without thinking about the form and function of particular objects. In this way, "Social structure is at once drawn on in human activities and at the same time (and potentially transformed) as a largely unintended consequence of those activities" (Faulkner & Runde, 2013, p. 13)

Faulkner and Runde use this way of conceptualising technological objects' identity by substituting an object's material form with its structure (consisting of constituent parts, their arrangements and interactions but not physical attributes). In this way, they argue, digital objects deserve close theoretical attention because of their special qualities. Thus, although artifacts are social and gain their identity from their social context, they have sustained identities and functions that can be examined at particular points in time.

77

Software represents new tools by which rhetorical conventions can be constructed. There are two key ways in which code has power according to Kitchin and Dodge (2011). Software is an actant with technicity; it does work in the world. Although code is invisible it 'produces visible and tangible effects' (p. 4). Software operates by codifying the world into 'rules, routines, algorithms, and captabases'. In this way, code is 'the manifestation of a system of thought – an expression of how the world can be captured, represented, processed, and modelled computationally' (p. 26) Software doesn't only represent the world but participates in it (Dourish, 2001).

The ideals that software reflects in its categorisation of the world become a reality when such categorisation becomes the common-sense way of framing particular social phenomena. When a census form enables respondents to choose only female or male categories, for example, this reinforces the gender binary. The ideal world as framed by the software reinforces a world categorized in this way. According to Kitchin and Dodge, one of the effects of the abstraction by software algorithms and data models in rendering aspects of the world is that 'the world starts to structure itself in the image of the capta and code – a self-fulfilling, recursive relationship develops.' (Kitchin & Dodge, 2011, p. 41)

The second point is that code has what Adrian Mackenzie (2006) calls 'secondary agency' in that it can be programmed to operate automatically without human oversight (Kitchin & Dodge, 2011, p. 5). An example here is the work of an algorithm that automatically parses data according to a set of encoded rules. Tarleton Gillespie (2010) writes that a class of algorithms that he calls 'public relevance algorithms' such as Google's search algorithms are a method of 'producing and certifying knowledge' and

should be subject to greater public debate since they represent 'a particular knowledge logic, built on specific presumptions about what knowledge is and how one should identify its most relevant components' (Gillespie, 2014, p. 168). The ways in which algorithms have this effect is in how they select what information is considered most relevant, manage our interactions by highlighting some contributions and excluding others, and reflect the public back to themselves, thereby shaping a public's sense of itself and with implications for who benefits from that representation.

According to Kitchin and Dodge, the power of code to execute its vision of the world in this way is dependent on human oversight or authorization. Technological systems are not entirely autonomous, but rather the autonomy of the system is a function of input, the sophistication of processing and the range of outputs that the code can produce.

Software may possess secondary agency but it does not have intent. For this reason, we need to look to those who design and deploy software have the potential to wield software for their own goals. David Berry (2014) writes that 'the new gatekeepers to the centres of knowledge in the information age are given by technologies, cognitive and data-processing algorithms, data visualization tools and high-tech companies' (Berry, 2014, p. 181). Algorithms and tools are seen to require technical 'black boxes' in order to simplify systems, but as a result they obscure how systems are constituted (Berry, 2014, p. 183). Black boxes consist of technologies that 'hide what is inside, sometimes productively, sometimes not, in order to simplify systems by hiding complexity or to create abstraction layers.' (Berry, 2014, p. 183) They are a defining feature of computationality, according to Berry.

The idea of technical obfuscation lies in stark contrast to the principle of transparency and openness that has become a key feature of the digital environment. Berry counters that the digital commodity is often available as an end with the means veiled and backgrounded.

> (T)he code(s) are themselves hidden behind an interface or surface which remains eminently readable, but completely inscrutable in its depths. (Berry, 2014, p. 197)

It is in this environment where code is obscured that the 'postmodern rich' will be able to have access to better cognitive support from computing rather than being better educated.

They will have the power to affect the system, to change the algorithms and even write their own code, whereas the dominated will be forced to use partial knowledge, incomplete data and commodified off-the-shelf algorithms which may paradoxically provide a glitch between appearance and reality such that the dominated will understand their condition in the spaces created by this mediation (Berry, 2014, p. 177).

Berry writes that truth is increasingly tied to expenditure and power, because the pursuit of knowledge is tied to the use of advanced (expensive) technologies and that this is already happening through the technological efforts to restructure the Web. In order to build this new computational world order, the existing gatekeepers of knowledge are already restructuring their data, information and media to enable the computational systems to scour the world's knowledge bases to prepare it for this new augmented age (p. 177).

As we will see in Chapter five and six, Wikipedia editing is almost entirely mediated by code. The travel of facts within this environment is, therefore, influenced by the

materiality of the code that produces the objects, tools, processes and conventions that drive Wikipedia. From the way in which editors profiles are presented and made searchable within Wikipedia's history pages, to the technical permissions of registered users to edit warning templates and the tools that have been developed to enable vandal fighting, the materiality of software code is vital to understanding how facts travel in Wikipedia. Code determines facts' travelling companions, their boundaries and terrain and even how the reproducibility of facts is governed. Along with discourse, ideology and boundary work, an analysis of coded objects and processes is vital to the framework that enables facts to travel on Wikipedia.

2.5 Co-production

Earlier in this chapter, I outlined the co-production idiom as a useful response to the weaknesses of some of the STS literature and as a practical framework for understanding the effects of socio-technical change on the material landscape. Jasanoff (2006) writes that, during time of significant social and technological change, traditional authorities are often reaffirmed, while at the same time there is a redefinition of identities as societies redraft social order in order to fill gaps in expertise that result from the adoption of new technologies. New power relations become lodged within identities, institutions, languages and representations.

> The identities, institutions, languages and representations created by science and technology can be politically sustaining, by helping societies to accommodate new knowledge and technological capabilities without tearing apart (by reaffirming) the legitimacy of existing social arrangements... When the world one knows is in disarray, redefining identities is a way of putting things back into familiar place. (Jasanoff, 2004, p. 39)

Technological and social changes, in other words, do not only involve the introduction of new roles, but also the re-affirmation of existing authorities. The identity of the expert,

in particular, is a core focus of co-productionist writings. The expert is a 'quintessential bridging figure of modernity' (Jasanoff, 2004, p. 39), arising out of the need to translate the growing complexities of science and technology for broader social goals.

In order to accommodate significant change, Jasanoff shows that there tends to be a re-drafting of the rules of social order regarding trustworthiness and authority. This redrafting is neither clear-cut nor does it happen immediately. Instead, boundaries between expert and non-expert, scientist and non-scientist are being continually reinscribed. In some cases the authority of certain experts is reaffirmed; in others, there is a translation of one type of expertise into the context of another as new expert identities are introduced.

Co-production is a useful framework for investigating the conditions under which facts travel and whereby authorities are being reconfigured. There are two clear benefits to the framework. Firstly, co-production avoids the charges of natural and social determinism because facts (emerging from science or other authorities) are neither a mirror of reality nor a result only of social and political interests.

Secondly, co-production responds to criticisms that STS is too internalist and lacks normativity to be critical (Jasanoff, 2004, p. 4). According to Jasanoff, co-production offers new ways of thinking about power, 'highlighting the often invisible role of knowledges, expertise, technical practices and material objects in shaping, sustaining, subverting or transforming relations of authority' (p. 4).

In order to understand how authority is being reconfigured within Wikipedia, however, there is a need to bring together the co-production idiom with theories relating to the

strategies, practices and discourses that define the knowledge environment, particularly in the realm of fact travel. Three core themes relating to Wikipedia's socio-technical system are explicated in the chapters that follow:

1. The traveling companions, boundaries and terrain, and character of facts;
2. The identities, discourses and representations in which power becomes lodged;
3. The social and technological sources of power and authority.

These three themes frame identities, discourses and representations within Wikipedia as the locations in which power becomes lodged and that are co-produced by social and technological forces. Identities, discourses and representations also influence the ways in which facts travel: identities determine who the authorities are and therefore which are considered 'good company'; discourses are the primary means of communities expressing and reproducing ideologies that constitute the terrain, and the characteristics of representations (or facts, in Wikipedia's case) can help ensure their travel or put them at a significant disadvantage.

The ways in which the landscape in which facts travel is constructed is through both discursive and performative means. Within the digital environment, performativity is a process of enacting digital speech conventions as well as defining actors according to their own ideological interests. As a result, power becomes lodged in identities, discourses and representations, with social and technological forces continually enacting the power relations that define them.

This does not, however, mean that the outcome of such networked relations is final, since other actors are continually joining networks and changing their dynamics. The

boundaries between science and non-science, professionals and amateurs, facts and theories are neither dualistic nor is their identity decided rapidly or once-and-for-all. Instead, such categories are enacted and performed in multiple locales in which particular identities and roles feature. In Lynch's (2006) terms, expertise and the associated domains of expert knowledge are a local interactional production rather than an expression of rules, policies or laws.

What becomes important, then, is in discovering the local battlefronts where territorial claims are being fought for and negotiated. Such battles play out in defining the roles and identities of knowledge actors. The tools used to define roles and identities are garnered from the socio-technical system in which knowledge claims are made. In Wikipedia's case, the socio-technical system is almost entirely mediated by code which means that we need to look to the particular affordances of that system, as well as how those affordances interact with users and practices, in order to fully understand the process by which Wikipedia and Wikipedians maintain their authority over factual information.

2.6   Conclusion

Can facts be shown to travel within Wikipedia's socio-technical system? If so, do some facts travel further and with greater integrity than others, and why is this so? Is there evidence that the roles of knowledge authorities and experts are being reconfigured within Wikipedia's socio-technical system, and if so, how has this reconfiguration taken place?

In order to answer these questions, this thesis uses theory from Science and Technology Studies relating to the ways in which knowledge communities construct the boundaries

between truth and fiction, science and ideology, empirical facts and theoretical conjecture. Central to this construction is the identity of the expert as a 'bridging figure of modernity' (Jasanoff, 2004, p. 39). Instead of new expertise being immediately recognized as technologies change, however, there is usually an intermediate process in which traditional authorities are reaffirmed while social order is reconfigured. Experts are recognized as such when they are able to speak the language of membership. In the digital environment, this language is both descriptive and performative, involving digital speech acts and conventions that need to be learned through the sharing of tacit knowledge.

In order to understand how expertise is being reconfigured within Wikipedia, there is a need to analyse the descriptive and performative language that is being spoken and enacted as facts travel through Wikipedia's socio-technical system. Facts travel well when accompanied by good traveling companions, a favorable environment and advantageous characteristics. The definitions of 'good', 'favorable' and 'advantageous', however, are dependent on the socio-technical context in which facts are traveling. The context that defines these terms can be found not only in what a knowledge community thinks it does, but also how those ideals play out in practice. In Gieryn's terms, we need to look at the ideals of what Wikipedians say they do but also what happens in the practice of particular facts and articles as well.

Using this framework, the next chapters provide an analysis of the role of each of these factors. Chapter four describes the boundaries and terrain of Wikipedia with particular emphasis on Wikipedia policies and norms. The principle of verifiability is discussed as foundational to Wikipedia's claims to authority and in Shapin's terms represents the literary, social and material technologies that support the travel of facts within

Wikipedia's socio-technical system. Wikipedia's policies are instructive of how it accords authority to different sources of knowledge, and what is required in order to become an effective speaker of the language of Wikipedians, but it is at the edges of the network, in Jasanoff's discourses and representations in which power is lodged.

Chapters five and six provide an analysis of two examples of the edges of Wikipedia's network in the discussion of two attempts to extend the boundaries of what facts Wikipedia will include in its corpus. Chapter five deals with the travel of facts about 'surr', a game played by people in the villages of northern India while chapter six covers the case of the 'Egyptian Revolution of 2011' article. In the case of surr, the attempt by the fact's traveling companions to propel it along proved unsuccessful; in the case of the Egyptian Revolution, the fact traveled rapidly and widely.

The success and failure of these facts is considered according to their traveling companions, the terrain and the character of the facts in order to two key questions answered in chapter seven. The first is what constitutes good company, favorable terrain and advantageous character. The second is what strategies are being used by Wikipedians in order to propel facts forward given their knowledge of the factors that determine success in this environment. The impact of Wikipedia on knowledge authority and expertise is discussed with an emphasis on which authorities Wikipedia is reaffirming and how expertise is being reconfigured. Finally, chapter eight discusses the implications of this for knowledge building communities outside of Wikipedia as well as areas for future research.

Chapter 3: Research design and methodology

This study constitutes a multi-sited ethnography of Wikipedia conducted over the past four years. I chose Wikipedia as the field site for this study for two key reasons. Firstly, in studying how new configurations of fact-building power are emerging in the context of the Internet, Wikipedia is becoming an important site of struggle, along with other collective fact-building platforms such as OpenStreetMaps and Freebase. Like OpenStreetMaps and Freebase, Wikipedia is a platform in which different actors are more or less collectively deciding on a particular narrative out of all the possible narratives that could be developed.

Secondly, Wikipedia is held up as an ideal, a manifestation both of the promise of the Internet to enable speech, but also, more recently, of the problems that occur when anyone can edit. Wikipedia is therefore a rich source of public discourse and debate that a wide spectrum of actors participate in, and whose representations have significant power in performing particular types of identities. The power relations performed by actors within the Wikipedia platform therefore have an impact on how similar platforms are designed and developed, especially due the growing centralisation of fact building in the age of big data.

It is important to note that I have chosen Wikipedia as a field site, rather than as a case. In defining the field site, I have selected a series of paths to follow at the same time as participating daily in a diverse range of Wikipedia practice. Whereas case selection implies the early decision about which groups of actors and documents to study, field selection implies the embodied traversing of different networks over time. Using the

metaphors of following (Marcus, 1995) I have engaged in the traversal of two cases of Wikipedia articles as well as traversing the foundational policy of verifiability within Wikipedia's socio-technical system.

I have employed ethnographic practices and principles in designing the methodology for this study. Ethnography is an approach to sociological research that applies 'to the acts both of observing directly the behavior of a social group and producing a written description thereof' (Marshall, 1994a). Jenna Burrell (2009) argues that ethnography is best understood not as a single method but as 'a complex of epistemological framings, methodological techniques, and writing practices (which) has spread into many domains and disciplines beyond its roots in cultural anthropology' (p. 181). These methodological techniques have at their center the role of the ethnographer or field worker who constructs their field site around their movements and encounters in experiencing life as a member of a particular community.

Ethnographic theory and practice has undergone repeated transformations in line with changes experienced by societies since the first ethnographies in the early twentieth century. In 1922, Bronislaw Malinowski published a unique study of the Trobriand people who live on a small island chain northeast of Papua New Guinea. His monograph entitled 'Argonauts of the Western Pacific: An account of native enterprise and adventure in the archipelagos of Melanesian New Guinea' dealt with themes of economics and social organization described through the trading system of the Trobriand Islanders. Along with similar studies that broke with past ways of documenting distant cultures, 'Argonauts' validated the persona of the fieldworker and led to greater emphasis on the power of observation and the role of theory in helping academic ethnographers to 'get to the heart' of a culture (Clifford, 1983). These

innovations served to validate ethnography based on scientific participant observation in the first part of the twentieth century (Clifford, 1983).

Since then, ethnography has evolved from the traditional view of the researcher traveling to a remote place in order to study a *whole* culture by learning its language and customs. Two of the major developments in ethnographic practice occurred as a response to the decline of the colonial project and the rise of global capitalism. These include the recognition that multiple sites are often required for any modern understanding of culture and the more recent acceptance of the Internet and virtual field sites as legitimate sites of study.

Christine Hine (2008) writes that ethnography as a method for studying online social formations emerged from the recognition that these were, in fact, communities rather than individuals making use of new technology. She points to Nancy Baym's work on Usenet soap opera fans (Baym, 1999) as one of the first studies to describe themselves as ethnographies of online communities where Baym argued for the social formation to be described as a community and that ethnography was an appropriate approach for studying   it.

There have been at least three previous ethnographic studies of Wikipedia. The first was 'Good Faith Collaboration' by Joseph Reagle in 2010, the second was an article entitled 'Janitors of knowledge: constructing knowledge in the everyday life of Wikipedia editors' by Olof Sundin (2011), and the third is the 2014 study by Dariusz Jemielniak entitled 'Common Knowledge? An Ethnography of Wikipedia'. Reagle's (2010) ethnography of Wikipedia focused predominantly on Wikipedia's policies and norms in trying to understand why Wikipedia has succeeded. Sundin (2011) focused more distinctly on

interviewing Wikipedia editors but ended up primarily interviewing administrators or those predominantly taking care of administrative tasks. Finally, Jemielniak's (2014) study is rich with ethnographic detail since the author became an administrator of the English and Polish Wikipedias, but this study, too, relied on interviews and experiences of administrators. Each of these ethnographies is a rich source of understanding about Wikipedia laws and norms, but each is filtered through the perspective of policy makers and administrators – that is, through the lens of those who already have significant power and experience on Wikipedia.

My ethnographic study aims at an alternative perspective, another frame with which to view what Wikipedia produces and what its place in our world is. In order to construct this frame, I have focused on ethnography's principle of multiplicity using two key techniques. Firstly, situating myself at the point of the article itself and following the actors involved enabled me to engage with a wide range of perspectives represented by the different actors that attach themselves to its success or failure. Additionally, I have attempted to seek out the actors and phenomena that are in the shadow of the encyclopaedia, those who have been banned, whose articles have been deleted and those who are opposed to what they believe is the hegemony of the mainly-Western, male Wikimedia Foundation and Wikimedia community. This goal has not been achieved without challenges (see below).

In the next section I outline how I approached the construction of the field site, participant observation, data collection and analysis, interviewing and memoing.

3.1 A multi-sited approach

This study constitutes a multi-sited ethnography because the fieldwork has consisted of traversing many different places and networks in order to explore different aspects of the phenomenon of fact building on Wikipedia. Much of the data collection was conducted online from Oxford in conversation with editors and other stakeholders around the world including Australia, Finland, Germany, Israel, South Africa, the United States, but also included data collection during Wikimania conferences in Washington D.C., Hong Kong and London, as well as during workshops with editors in Cairo, Amman and London. National boundaries were not the only ones traversed during this study. Although I focused on Wikipedia, I extended data collection to other Wikimedia sites including Wikimedia Commons, Wikidata and to sites outside of the Wikimedia group to Google, the New York Times, Al Jazeera and other sources.

In the construction of my field site, I have been guided by George Marcus's (1995) 'Ethnography in/of the World System'. Marcus argues that there is a trend in ethnographic research in moving out of single sites and local situations to look at the circulation of cultural meanings, objects and identities and that this is the result of both the ideas and concepts of postmodernism but also about empirical changes in the world and transformed locations of cultural production.

Marcus outlines the features of multi-sited research as being designed around 'chains, paths, threads, conjunctions, or juxtapositions of locations in which the ethnographer establishes some form of literal, physical presence, with an explicit, posited logic of association or connection among sites that in fact defines the argument of the ethnography' (Marcus, 1995, p. 105). Such an association between sites could be practically brought into being, argues Marcus, using tracing mechanisms for defining the

fieldsite along a malleable path that is extended as it is followed. These tracing mechanisms included the ethnographer 'following the people', 'following the thing', and 'following the metaphor', amongst others. In so doing, Marcus highlights how movement is central to social practice and how cultural processes can take place across great distances and on the move (Burrell, 2009). Marcus acknowledges that the ethnographer would need to renegotiate identities in different sites and that the result would be the portrayal of a 'slice of the world system' rather than a description of a 'whole culture' previously produced (Marcus, 1995, p. 105).

'Following the thing' as a method is the antithesis of the traditional ways in which fieldsites are designed prior to the project where the researcher would go to the place they had chosen to study, examining people, places and things within it and going away to do the analysis. In contrast, the fieldsite is determined by the researcher 'follow(ing) the thing' through the code/spaces that it takes them. Ethnography is about local, close-up perspectives - therefore we can use that to discover new paths of connection and association but on a differently configured spatial canvas (Marcus, 1995, p98).

Ethnography has played a significant role in the field of science and technology studies. According to Hess (2001), the laboratory ethnography constituted the first stage of science studies research, followed by more general studies when there was greater interest in issues of culture and power that made researchers move outside the laboratory. Laboratory studies enabled ethnographers to examine the practices which constituted scientific facts and to engage with philosophical statements about the nature of science but the problem was that they didn't deal with the social forces beyond the laboratory walls (Hine, 2007, p. 659).

According to Golinski (1998, p. 11), Latour and Woolgar's study claimed that 'interactions among small groups of researchers were no less "social" than large-scale forces, such as classes or political movements' but this was much more a restricted specification of the social context relevant to understanding scientific practice than had previously been asserted. Hein (2007) argues that multi-sited ethnographies are appropriate for extending studies of scientific and technological practice beyond the walls of the laboratory. Ethnographies are based on the fundamental principles that 'it is in the local that the global emerges, indeed that there is no supra-local phenomenon except in so far as it is constituted in the local' (Hine, 2007, p. 656). Traditional spatial boundaries are an artificial means of construing the objects of research, particularly when they involve culture (Amit, 2000) and in a contemporary society that is increasingly characterized by mobility, connection and communication (Urry, 2000).

Multi-sited ethnography is one way of recognizing the world as complex and networked, that phenomena are experienced in multiple sites and are affected by multiple social and technical variables. According to Hine, multi-sited ethnography is a 'a faithful reflection of lives lived not in discrete locations, but through various forms of connection and circulation.' (Hine, 2007, p. 656)

Burrell (2009) provides practical methods of constructing the field site in the context of online social life using Marcus's multi-sited ethnography as inspiration. Burrell conceived of her field site as 'a network composed of fixed and moving points including spaces, people, and objects' (p. 189).

> (An) advantage of defining the field site as a network is that it is produced as a continuous space that does not presume proximity or even spatiality in a physical sense. Continuity does not imply homogeneity or unity; it implies connection. The continuity of a network is evident in the way that one point can (through one or more steps) connect to any other point. (Burrell, 2009, p. 190)

In a 'field site as network, the point of origin, the destination(s), the space between, and what moves or is carried along these paths is of interest to the researcher (Burrell, 2009, p.190). The field site is comprised, not only of people but of objects as well. According to Burrell, 'The field site as heterogeneous network incorporates mapping out the social relations of research participants and their connections to material and digital objects and physical sites.' (p. 191) This is important because some networks are defined by material flows rather than just relationships.

The focus of this thesis, then, is on culture as locally constituted, where culture is defined, not as 'a power, something to which social events, behaviours, institutions, or processes can be causally attributed; it is a context, something within which they can be intelligibly - that is, thickly - described' (Geertz, 1973, p. 316). What is required is densely packed facts, a 'thick description' of phenomena, rather than randomly sampling the population in order to discover universal truths.

The process of producing thick description within the field of ethnography is developed by Geertz (1973) from the writings of Gilbert Ryle (1968). Geertz describes ethnography as 'an elaborate venture in... thick description' (p. 6). Ethnography, according to Geertz, is essentially about interpreting social behavior by 'sorting out the structure of significance' (p. 312). Ethnography is less about decoding what Ryle called 'established codes' but rather acting more like a literary critic.

> The ethnographer "inscribes" social discourse; he writes it down. In so doing, he turns it from a passing event, which exists only in its own moment of occurrence, into an account, which exists in its inscriptions and can be reconsulted (Geertz, 1973, p. 312)

According to Geertz, the key characteristics of thick description are that description is interpretive of the flow of social discourse and that it is microscopic.

> The aim is to draw large conclusions from small, but very densely textured facts; to support broad assertions about the role of culture in the construction of collective life by engaging them exactly with complex specifics. (Geertz, 1973, p. 321)

Thus, the goal of case selection is not to be able to generalize across cases but rather to generalize within them (p. 320).

3.2 Case selection

I was continuously looking for examples and experiences that challenged what were held up as ideal, or widely accepted forms of knowledge and sources in my fieldwork. I did this in order to understand how the boundaries of what Wikipedia represents are being articulated through daily practice. I chose the article as a bounded case because there were clear beginning points that could be traced and mapped over time, and with Wikipedia's provision of multiple lenses through which to study the progress of the article over time, as well as the actors involved, provided a rich source of data to ground the studies in.

I chose the two cases according to a theoretical sampling method in which I focused on articles that fulfilled a number of criteria that would enable them to be illustrative of boundary work occurring on Wikipedia's outer edges. Firstly, I was looking for cases that had a significant political impact. Although all articles on Wikipedia are political because they are concerned with the representation of knowledge, I wanted to choose cases that had a high degree of political significance in the sense that their representation could materially affect the lives of their subjects. Secondly, I wanted to use articles that reflected diverse claims by numerous and diverse sets of actors (diverse in the sense of their goals and experiences with Wikipedia). For this reason, articles that were somehow in the news or subjected to wider debate beyond the article

and its talk pages were particularly useful. Finally, I wanted to look at cases where there were differing levels of success in articulating new boundaries.

I chose the articles, 'surr' and 'Egyptian Revolution of 2011' as unsuccessful and successful attempts to broaden the boundary of what Wikipedia represents in its corpus. In the surr case, actors who were calling for Wikipedia to include oral citations as new sources of knowledge within Wikipedia were largely unsuccessful in their claim making, and, as such, surr remains a relatively thin description or representation of its subject on Wikipedia. In the Egyptian Revolution of 2011 case, editors who wanted Wikipedia to represent breaking news topics were successful in their efforts to represent the protests of 2011 and so to perform Wikipedia as a source for information about breaking news topics. Both of these articles are significant in the political effects. Surr's success would serve as an example of how Wikipedia might counter challenges in the availability of sources about a large proportion of the world's known phenomena; its failure would serve as an example of how Wikipedia is being shaped towards the needs and experiences of a particular group of individuals at the expense of others. Finally, both articles became subject of (greater or lesser) traditional and social media attention.

There is a high degree of variability in the cases. The Egyptian Revolution of 2011 article was never nominated for deletion and enjoyed high levels of editing activity and readership and in 2011, it was one of the most viewed and most edited articles on Wikipedia (Keegan, Gergle & Contractor, 2013, p.2). Surr, on the other hand, had a significantly smaller edit count and article viewership, constituting a part of the long tail of Wikipedia articles in terms of popularity (see table 3.1).

Table 3.1 Comparing the traveling companions of surr and the 2011 Egyptian Revolution on English Wikipedia

Traveling companions	Surr	Egyptian Revolution of 2011
Number of editors	12	1,994
Number of edits	21	7,988
Number of other language versions article appears in	2	57
Article viewer numbers*	526	366,635
Number of page watchers	N/A**	260
Number of redirects***	0	56
Links to the page within Wikipedia	4	972
Links from the page	5	953

Notes:

* number of times the article has been viewed in the first three months of its creation (data sources: *http://stats.grok.se/* for article traffic statistics)

** fewer than 30 therefore not represented by the software at *http://tools.wmflabs.org/xtools-articleinfo*

*** redirects are alternative article titles (for example, 'Jade Revolution' and '25 January Revolution' that are automatically redirected to the main '2011 Egyptian Revolution' article title (see *http://dispenser.homenet.org/~dispenser/cgibin/rdcheck.py?page=Egyptian_Revolution_of_2011* for a list of redirects)

Source: http://tools.wmflabs.org/xtools-articleinfo, http://stats.grok.se/ as of 27 April 2015

The two cases are analysed against the background of historical analysis of other encyclopaedias as compared to Wikipedia's styles, policies and norms. Analysing printed encyclopaedias, I discovered a key difference between Wikipedia in terms of stylistic presentation of authorship information. This led me to discover the origins of Wikipedia's authorship styles as originating in multiple documents relating to policies, guidelines and essays, as well as interviews with founder members of Wikipedia's community and discussions on media sites relating to founding values and principles of

Wikipedia and related sites and movements. Table 3.2 illustrates the key data sources for both the case studies and the historical analysis.

For each of the case studies, I followed the articles and the actors, collecting data from interviews and a wide variety of documents and sources. I defined actors and actants according to their role in the constitution of the phenomena being represented in the article. Thus, in addition to the editors who edited the article, the templates that they used to frame the content as well as other actors who commented on the article on other platforms, I also included the subjects of the article themselves. Although I wasn't able to interview these subjects, they formed an important part of the analysis, thus avoiding the problem with SCOT's conception of the 'relevant social actors' noted by MacKenzie and Wajcman (1985, p. 22) as only those actors who are explicitly associated with the artefact or claim.

Table 3.2 Data and documents analysed

Historical analysis of Wikipedia's values, principles and encyclopedic identity

Policies, guidelines, essays	Wikipedia's core content policies (verifiability, Neutral Point of View, No Original Research); Wikipedia's 5 pillars; analysis of essays relating to verifiability
Encyclopaedias	Encyclopaedia Britannica (from 1768); Penny Cyclopedia (1828 to 1843); Tomlinson's Cyclopaedia of Useful Arts (1852-1854); Cyclopedia of General History (1880-1884)

Surr case

Actors	Oral citations (interviewees, interviewers, authors), editors, Wikimedia Foundation
Interview data	Achal Prabhala (2 hours + follow up emails), Utcursch (email interviews), Fifelfoo (IRC interview), Joseve05a (IRC interviews), Jimmy Wales (in person interview)
Documents	Policy documents, talk pages, Reliable Sources noticeboard discussions, Wikimedia-l mailing list thread on oral citations, project page, edits to surr in English, Korean and Hindi
Media	Film, 'People are Knowledge' produced by Achal Prabala (see https://vimeo.com/26469276) Articles in the traditional news media including the New York Times and The Hindi (4) as well as 22 blog posts about the project listed on https://meta.wikimedia.org/wiki/Research:Oral_Citations

Egyptian Revolution of 2011 case

Actors	Article editors, subjects of the article, sources, bots
Interviews	Interviews with [The Egyptian Liberal], [Ocaasi], [Aude], comprising the core group of editors of the 2011 Egyptian Revolution article (totalling about 4 hours)
Documents	Talk pages (about 150 pages), policy documents, tools documentation, editor profiles, edits to the article in English

3.3   Methods and principles

Participant observation is a research technique which 'aims to gain a close and intimate familiarity with a given area of study through an intensive involvement with people in their natural environment' (Marshall, 1994). Participant observation involves a variety of methods including informal interviews, direct observation, participation in the life of the group, collective discussions, analyses of documents produced within the group, self-analysis and life histories (Marshall, 1994).

One of the key strengths of participant observation is that it provides opportunities for embodied learning by the researcher. By participating in the practices engaged by particular communities, ethnographers gain a detailed, embodied perspective of social situations in context. As Larsen notes,

> Compared to qualitative interviews, observations better capture the bodily, enacted, technologised and "here-and-now" quality of practices because they focus on immediate physical doings and interactions rather than retrospective and reflexive talk about how and why such performances take place, and what they mean. (Larsen, 2008, p. 13)

The potential weakness of participant observation is that the observer's role may not provide access to the total population under investigation. As an infrequent editor, I largely encountered Wikipedia as an editor with few edits to my username so that I could experience the Wikipedia world from the perspective of a user on the edges of the Wikipedia community. I supplemented this perspective by interviewing individuals who occupy a variety of roles within Wikimedia, and I performed extensive document analysis to understand editing and governance practices.

In the initial research phase, I constructed my field site in order to receive daily updates of community activities from a variety of sources. These included issues relating to Wikipedia politics, power and authority on a number of mailing lists, on Twitter,

Facebook and on a number of different news sites (see table 3.3 below). My days would begin by checking email where I was signed up to multiple mailing lists in which members of the Wikimedia community would discuss and debate issues of important. I would then check my Twitter feeds where I was signed up to receive Tweets relating to keywords such as "Wikipedia + deleted" so that I could understand challenges feed by actors outside of the Wikipedia platform as well as investigating how internal debates were being covered both by traditional media and within social media platforms. I also checked updates from relevant Facebook groups (listed below), including discussions relating to controversies that were being discussed in by the media and within academic fields.
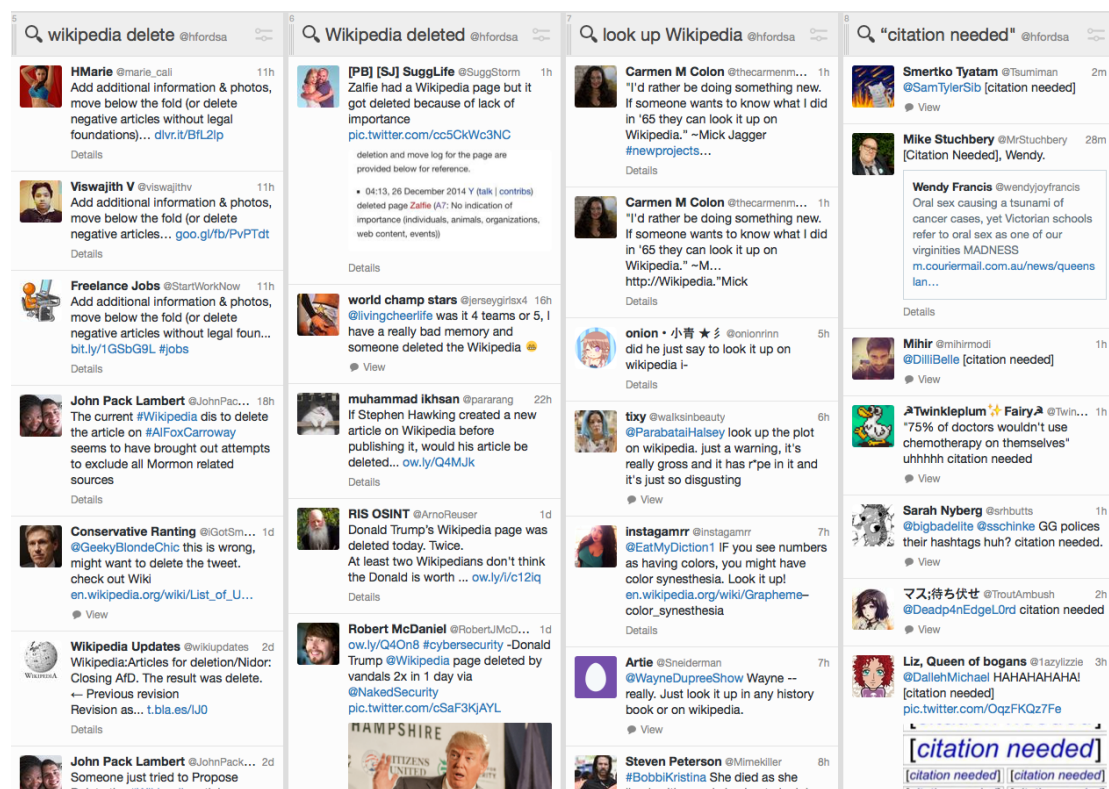
Figure 3.1 Screenshot of Twitter feeds for four keyword searches

Source: Twitter.com, 27 July 2015

These daily updates constituted a daily awareness filter for noticing what was being spoken about in the community as well as how Wikipedia was being responded to by its

editors, by WMF staff, by the media and the larger social media publics. As Christine

Hine writes,

> Even an asynchronous message board has its own version of "real time", as the
> rhythm of changing activity levels across days, weeks and months, and the speed
> with which messages receive replies, are parts of the experience that members
> have of their group. An ethnographer who simply collected a corpus of messages
> in one visit would miss that experience. (Hine, 2008)

Table 3.3 presents a list of data sources consulted in my daily field experience.

Table 3.3 Table of key sources from the field

Media/forums	Sources
Mailing lists	Wikimedia-l: Wikimania general list Wikidata: Discussion list for the Wikidata project Commons-l: Wikimedia Commons Discussion List Wiki-research-l: Research into Wikimedia content and communities Languages: Languages discussion and Wikimedia Indigenous Languages
Online news	Wikimedia Signpost: Weekly community-authored newspaper covering issues relating to the English Wikipedia and related projects as well as the Wikimedia Foundation Wikipediocracy:
Twitter	Following Wikipedia users, Wikimedia Foundation accounts Feeds on keywords: "delete + Wikipedia", "Wikipedia + deleted", "look up Wikipedia", "citation needed" (see figure 3.1 below)
Facebook groups	Wikipedia Weekly: Wikipedia in the news Wikipedia Women: Closed group for discussions among female Wikipedia editors menaWikipedia: Discussions relating to Wikipedia in the Middle East and North Africa Wikipedia Wikipedia Tunisia: Discussions relating to Wikipedia in Tunisia Wikipedia Arab Community: Discussions relating to Wikipedia in the Arab world
Quora	Following Wikipedia users and topics relating to Wikipedia

In the second phase, I participated in a number of editing activities, including writing

articles and defending them against deletion on non-notability grounds on English

Wikipedia, working with a seasoned vandalism fighter to delete vandalism using the tool and participating in discussions on mailing lists. In January 2013, I produced an article for English Wikipedia on the South African internet meme, '(Don't) touch me on my studio' (https://en.wikipedia.org/wiki/(Don%27t)_touch_me_on_my_studio) which had resulted in numerous video remixes and jokes on social media. Writing this article using the article creation tool enabled me to gain first-hand experience of both the process and emotions associated with article creation and evaluation. The article was nominated for deletion about 45 minutes after it was first published using the tag:

> WP:NOTNEWSPAPER. Doesn't adequately demonstrate a lasting significance
> and encyclopedic value (from
> https://en.wikipedia.org/w/index.php?title=(Don%27t)_touch_me_on_my_studi
> o&oldid=535129355)

I sought assistance on social media in order to oppose deletion, as well as improving the article and responding to criticism on my talk page. The proposed deletion tag was eventually removed but only because I had spent considerable time learning how the process worked, following the profiles of concerned editors and crafting responses that fit Wikipedia's deletion process vocabulary. Experiencing the process of deletion myself was critical to understanding the perspective of those whose content has been questioned or removed on Wikipedia over time.

I also acted (briefly) as a vandal fighter so that I could obtain the perspective of someone on the evaluation end of new content creation. After interviewing Joseve05a who had removed vandalism from the Egyptian Revolution of 2011 article, I was invited to use the Huggle anti-vandalism software. Mentored by my [Joseve05a] [15] who talked me through the process on IRC, I encountered live edits being filtered through Huggle

[15] Wikipedia users are indicated in the text using [ ]s.

software and mistakenly deleted an edit that was actually made in good faith. My mistake was not without consequences. [Joseve05a] had to explain to the administrator who had granted me access to the software that I would be more careful in the future, and realized how easy it was to make a mistake that left very thin traces for the user whose content had been deleted to piece together. This experience was extremely useful for me in understanding how detached many users were from the content of the article itself, since their editing was mostly performed remotely using standalone software or bots.

I also participated in a number of offline workshops for Wikipedia editors and interested stakeholders including Wikimania in Hong Kong in 2013 and in London in 2014, I have also been involved in Wikipedia's research community, being part of the planning committee in 2013 of WikiSym and occasionally helping to edit the 'Wikipedia Research Newsletter'. Talking to the Wikimedia Foundation about my research results and their implications for Wikimedia Foundation policy and strategic practice, particularly in the developing world.

Studying social interaction online has the benefit of significant amounts of traces of such interaction available for the researcher to make use of. Wikipedia, for example, is a potentially rich area for studying information generation practices because not only do we have access to discussions that serve to coordinate the editing of a rapidly-evolving article (in the form of talk pages and others discussion areas on Wikipedia); we also have access to successive versions of the article as it endures hundreds – sometimes thousands – of edits (in the form of the article history). Wikipedia data enables us to trace detailed activity around particular editorial decisions because of the existence of

time and authorial metadata. Access to these discussions affords a unique opportunity for researchers to understand the practices that give rise to particular representations.

The availability of Wikipedia's comprehensive historical records enable the researcher to reconstruct the events and the conversations that led to particular decisions being made and so are a potentially incredibly valid way of researching. By reconstructing each edit over time and drawing from relevant conversations having a potential impact on such edits, the ethnographer can draw together the actions and their related discourse in order to understand the social milieu in which editing decisions were being made. Editors themselves undergo similar reconstructions when they orient themselves in Wikipedia's spaces, following conversations on the talk page before they edit the article and reading history pages to see who the frequent editors are and who they are reverting, for example. Using time stamps, user profiles and links to other locations on the web and Wikipedia, researchers are able to recreate this rich social setting, undertaking what Geiger and Ribes (2011) call 'trace ethnography'.

> Analysis of these detailed and heterogeneous data – which include transaction logs, version histories, institutional records, conversation transcripts, and source code – can provide rich qualitative insight into the interactions of users, allowing us to retroactively reconstruct specific actions at a fine level of granularity. Once decoded, sets of such documentary traces can then be assembled into rich narratives of interaction, allowing researchers to carefully follow coordination practices, information flows, situated routines, and other social and organizational phenomena across a variety of scales. (Geiger & Ribes, 2011, p. 1)

Trace ethnography is certainly a helpful way of understanding how these rich details can constitute a script of what happened by whom and when, but it still leaves out the participation feature that offers an opportunity for the researcher to engage in embodied learning and gain from the deeper levels of understanding that come with being personally engaged in the experience under study. Larsen argues that ethnographic observations must be supplemented by qualitative interviews since

'interviews that allow space for the unexpected and people's accounts of how their performances are meaningful are vital to avoid portraying 'performers' as 'cultural dopes'. (Larsen, 2008)

Interviews are useful for discovering data missing from such traces including the editor's emotional state when editing or when confronted with particular communicative responses, the logic of individual editors when making editing and communication decisions, conversations between editors on private spaces that are not tracked, and the high and low points for editors in when reflecting on their experience in the past. In this study, interviews were essential for recognising that [The Egyptian Liberal]'s Wikipedia account was being shared among three other editors in different countries, for hearing about the biases that [Ocaasi] admitted to having against and then for Al Jazeera before and during the 2011 Egyptian Revolution, and well as [Ocaasi]'s reflections on particular edits that had a significant impact on him. This rich information could only have been garnered through in-depth qualitative interviews.

I used different sampling techniques for the two cases that I analysed, depending on the numbers of actors who were involved in the construction of the article. In the surr case, I was able to interview almost all the editors of the article because of the low numbers of editors involved in editing. I invited each of the editors of the surr article to an interview. Four out of the six answered my request and I conducted several rounds of interviews with each of them. In the Egyptian Revolution case I chose the most frequent editors.

I interviewed on Skype, email, IRC (Internet Relay Chat) and Google Hangouts depending on the preferences of interviewees but also conducted some interviews

during face-to-face gatherings. I formulated a method of doing interviews with informants that I have (with Elizabeth Dubois) called 'trace interviewing' (Dubois and Ford, 2015) that draws from ethnographic approaches to specifically use trace data in the conduct of interviews in order to reconstruct the activities that users were engaged with. This method was helpful in connecting digital trace data with the personal analysis and narratives discussed by editors as they explained traces left of their experience in the data.

In the case of surr, I analysed each edit, following trails of article summaries to related discussions on various forums (including the reliable sources noticeboard and the Wikimedia-l mailing list). I also conducted interviews with three members of the Wikimedia Foundation who represent management, software development and executive levels of the organisation. Other documents consulted included project pages, user profiles, policy pages and software related pages.
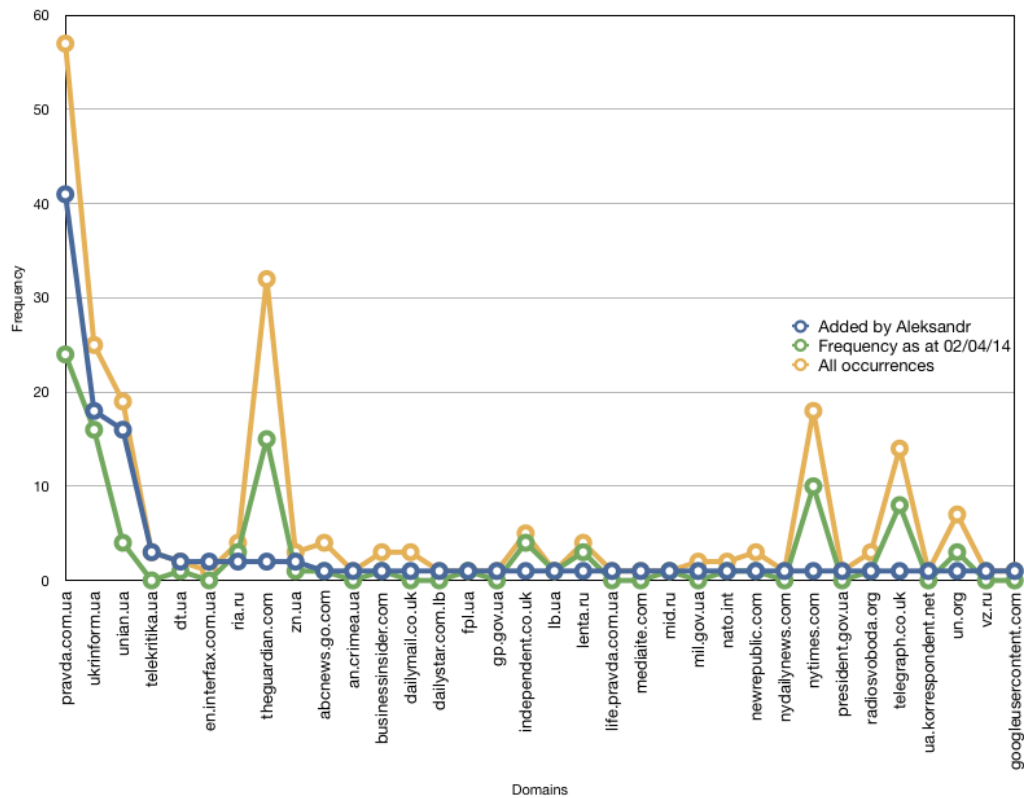
Figure 3.2 Graph used in one of the interviews for trace interviewing

Source: author

Trace interviews involve the collection, visualization, and discussion of a participant's traces with that participant. Visualized data of a user's interactions are employed in the interview setting so that the participant might reflect on her actions as depicted in the data (see figure 3.2). This process enables participants to interpret data by providing contextual details and clues about their motivations for undertaking particular actions represented in the data as well as to point to missing or inaccurate data. During the interview, the participant is asked to examine and then comment on the visualizations by answering questions to guide interpretation, such as, "Is this [data] surprising?" and "Can you explain X [data point]?"

Finally, memoing was something I did regularly throughout the process where I documented my thoughts about the data that I had encountered, the conversations that I

108

had had, and exploring my ideas for moving forward. In the intense data collection stages, I wrote memos as daily routine practice, summarizing interviews and my reflections about them, linking to emerging research themes and highlighting next steps. I also wrote thematic memos relating to particular documents together and used those as a basis for writing chapters on different concepts or issues. Figure 3.3 below indicates an example of a memo that I wrote after reading about a Wikipedia controversy that received a significant amount of press attention in 2014. I used this as inspiration for the analysis of the strategic wrangling going on behind the scenes of Wikipedia articles.
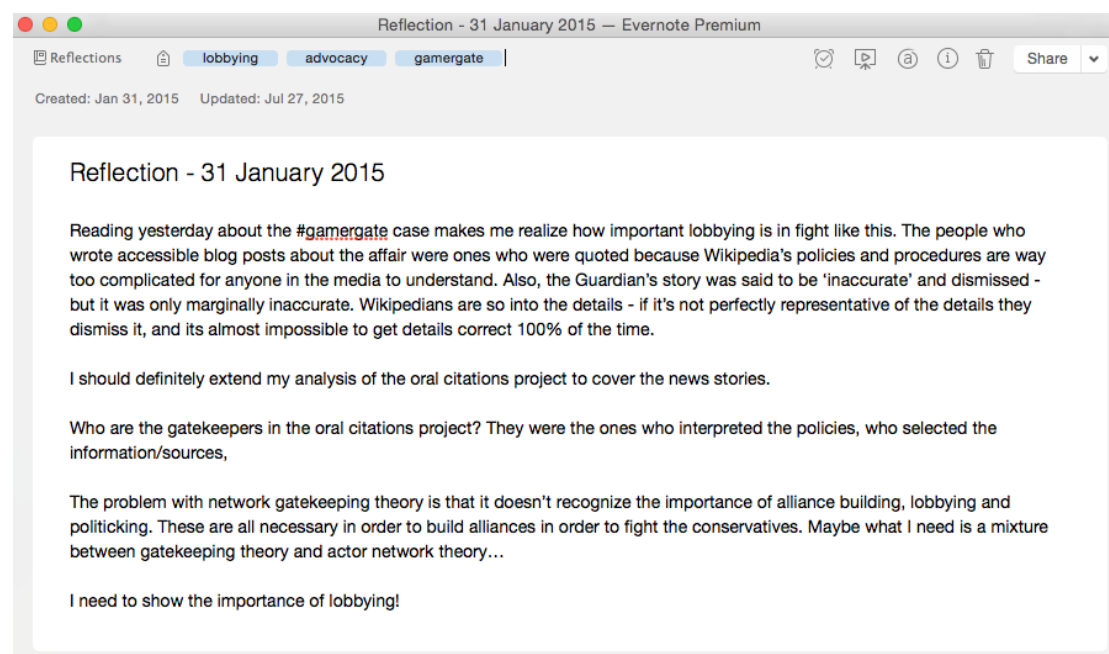
Figure 3.3 Memo from 31 January 2015

Source: author

3.4 Data analysis

All documents were coded using NVivo according to categories relating to skills, expertise, authority, credibility, experience, knowledge, as well as according to the actors' activities, membership categories and strategies. Analysis proceeded by first

building a narrative of what happened at what points in time and space, and then analysing the data according to themes relating to the research questions.

I used two methods of data analysis: qualitative coding of interviews and documents using a coding scheme that was guided by the theory and research questions as well as the coding of sources and citations in the Egyptian Revolution article. Tools were coded according to affordances and features. I performed qualitative coding using a process of iterative classification of data according to numerous heuristics, re-coding as I developed a deeper understanding and recognition of patterns in the data and then moving to a comparison of categories as I progressed towards final analysis.

Because of the large disparity in numbers of sources added to both articles over time, I had to find ways of efficiently analysing the almost 1,600 citations that were added and removed from the Egyptian Revolution of 2011 article over time. I used a few different lenses with which to describe this data, employing classification mechanisms tied to my theoretical framework.

I first downloaded all citations added to the article from the first edit on the 25th of January 2011 to the time of data collection on the 14th of July 2014. This consisted of a list of 1,565 citations from 401 separate domains or sources. 97% were represented by URLs and 3% were citations to books that were not available online.

I analysed the types of source actors represented at the time of data collection according to categories of sources and compared this with categories over time in order to understand the dynamics that lead to the stability or instability of sources over time. I categorised sources according to creator type and drilled down into the category of

academic sources to classify sources according to the document type, as well as according to where the source occurred within the article.

I then developed a codebook for creator categories based primarily on how authorship is attributed, whether to the individual or group, and then according to the individual's relationship to the organisation and the type of group or organisation (see table 3.4). The spectrum of unaffiliated to affiliated/group authorship ranged from, on the one hand, unaffiliated individuals, to academics and journalists who are explicitly affiliated to particular organisations, then civil society organisations and governments where authorship is attributed to the group rather than the individual.

Table 3.4 Creator categories codebook

Category	Definition
Unaffiliated	Individuals whose affiliation is not directly mentioned (or obvious) in the authorship of the content
Academics	Authors who are formally affiliated to universities and other educational and research institutions
Journalists	Media workers who are formally affiliated to a traditional media organisation
Organisations	Organisations such as Human Rights Watch who attribute reports and documents to the organisation rather than named individuals; also media companies when no individual author is named
Governments	Documents attributed to governments (predominantly from .gov domains)

Furthermore, I categorised sources in the academic category in order to understand which of these sources were being performed as authoritative within Wikipedia's representation of the Egyptian Revolution. In order to do this, I classified sources according to their main media type and located the citations in either the main text of

the article or in the 'external sources' links. In doing so, I was able to understand the particular placement of sources within the article.

3.5   Navigating challenges

The challenges of multi-sited ethnography include the heavy reliance on being able to gather the right kinds of data, and the risk of reflecting only a partial perspective. Multi-sited ethnography relies on the field worker's conscientiousness in gathering the right kinds of data (data that does, in fact, connect in consistent ways) and in being able to single out a particular narrative from the massive amount of data collected. The challenge in the multi-sited ethnography is not where one begins the research but in whether one exhausts all avenues in the tracing process. If only certain groups of editors are interviewed, for example, then the benefits of multiplicity and diversity would be lost. For example, if I had only interviewed advocates of the Oral Citations Project, I would not have learned about the strains faced by opponents to the project, nor that there were a variety of perspectives, many of them based on the implications of oral citations for practical rather than ideological reasons. It was only by seeking out alternative positions that I was able to recognize the role of conventions in limiting what can be represented on Wikipedia.

The vast amounts of diverging perspectives that are gathered through the ethnographic process can result in a cacophony of data when the research is written up. Focusing on how to reconcile these tensions in a clear narrative is therefore critically important for the researcher. Furthermore, Clifford (1983) notes that it is important to recognize that ethnographies are written constructions and thus subjective and that any experience enjoyed by the researcher in the social subject under study will be partial. 'The actuality of discursive situations and individual interlocutors is filtered out' in the published text

and the researcher effectively disappears into his text as an omnipresent spokesman (Clifford, 1983, p.132). Imagining that the researcher is directly conveying the truth of the social situation to others is unrealistic, in other words. Through language and writing and the researcher traditionally 'going away' from the field site to 'write up' their experience requires a particular interpretation that leaves the original experience in the dust of the final account. This is not to say that participant observation is not an essential part of the ethnographic process but only to recognize how experience of being in the field can be both varied and virtual.

In order to overcome these challenges, the research must continuously reflect on the extent to which they are exhausting all avenues in the tracing of activities or objects and that, in the presentation of results, that they continually refine their argument to reflect on theory but to do so as grounded in data. In this study, I have attempted to gain access to the explanations and experiences of as many of the parties representing the subject as possible and then to gather a narrative together that focuses on particular elements presented by the data and informed by theoretical perspective. The result is a single, clear narrative (chosen from multiple possible narratives) that is clearly grounded in the vocabulary used by actors to describe their experience.

3.6   Ensuring quality in qualitative enquiry

Four core concepts related to quality in qualitative enquiry as well as ethnographies in STS in particular are relevant to this study. They include the principles of crystallization, multivocality, complexity and the ability to intervene. For quantitative research, credibility is earned through reliability, replicability, consistency, and accuracy (Golafshani, 2003). Qualitative research, however, is primarily about using the

researcher as research instrument and therefore different criteria must be used for ensuring quality in qualitative research (Tracy, 2010).

Instead of the concept of triangulation which assumes a single point of view around which different data sources orient themselves in the realist paradigm, the concept of crystallization is a more useful metaphor for trying to understand how multiple types of data and methods are useful in 'open(ing) up a more complex, in-depth, but still thoroughly partial, understanding of the issue' (Tracy, 2010, p. 844). Closely related to crystallization is the principle of member checking where the researcher seeks input during the analysis. Tracy calls these 'member reflections', opportunity that 'allow for sharing and dialoguing with participants about the study's findings, and providing opportunities for questions, critique, feedback, affirmation, and even collaboration' (Tracy, 2010, p. 844).

In this study, I have provided opportunities for interviewees to comment on early drafts of case studies and in presenting drafts of research to Wikimedia community through presenting initial thinking at Wikimania conferences in 2012 and 2014, and during online workshops streamed by the Wikimedia Foundation in early 2015. On a number of occasions, member checking has enabled me to hone my interpretations of social behavior within particular contexts that I have interpreted, and to present my research in ways that are relevant to the community and the questions that I have been asking.

Multivocality is a principle about the discovery of diverse points of view in qualitative analysis. According to Tracy, multivocality emerges when social action is analysed from a participant's point of view and according to local context in order to interpret meaning (Tracy, 2010, p. 844). Multivocality requires the ethnographer seeking out viewpoints that diverge with those of the majority or the author. According to Hess (2001) who has

analysed ethnographies in the field of STS over the past few decades, ethnography, in particular encourages this approach and that the complexity that necessarily arises is a signal of quality work in the field.

> The methodological approaches which respond to and reinforce the experience of complexity include: the tendency to range around in the pursuit of appropriately engaged studies of our research objects; a focus on diversity as the key insight of ethnography, resisting the urge to find common threads or singular stories; and a tendency to stretch the notion of ethnography beyond comfortable limits. (Hess, 2001, p. 668)

I have applied the principle of multivocality to the case studies and to my analysis of practice more generally by constantly seeking out individuals who represent different points of view and represent different stages of Wikipedia community membership about particular topics. I have discovered a variety of different voices by focusing on controversial areas of Wikipedia editing and by using the SCOT approach to identifying and understanding 'relevant social groups' to a topic, but there are some challenges in being able to interview actors who are vandals or once-off editors since there is no practical means of connecting with them. Generally, however, the principle of multivocality has enabled me to gain a richer, more complex perspective on the multiplicity of goals, worldviews, epistemologies and experiences that Wikipedia editing involves.

Finally, both Hess and Tracy note the importance of ethnography in the field of STS and qualitative research more generally to intervene in the world that is described. According to Hess, ethnography in STS research is about 'both shaping a picture of the way the world is, and making propositions about its own role in accounting, and intervening in that world' (Hess, 2001, p. 668) and Tracy notes that a significant criteria for quality in qualitative research is whether the knowledge is useful.

> Practically significant research asks whether the knowledge is useful. Does it help to shed light on or helpfully frame a contemporary problem? Does it empower participants to see the world in another way? Does it provide a story that may liberate individuals from injustice? (Tracy, 2010, p. 846)

In April 2015 I participated in the Wikimedia Foundation's Research Showcase[16] where Wikimedia researchers are asked to discuss their research during a streamed video session. I presented the initial results from chapter four of my thesis framed according to a question of what actions the Wikimedia Foundation could possibly take in terms of the Oral Citations Project in order to enable the inclusion of content from the developing world. Focusing on the role of different actors and their power relations in shaping what Wikipedia represents is a topic that is relevant to the questions that the Wikimedia Foundation is currently asking itself, and as such I believe that this research will be useful as another lens with which to understand the cultural problems that have arisen in Wikipedia that are due to a number of complex factors.

3.7   Ethical considerations

Tracy (2010) notes that ethical research is an important goal for qualitative research and that such ethical procedures are contextual and arise out of the particular arrangement of the researcher and her subjects. In ensuring the ethical nature of this study I have taken two important steps: the first is in completing necessary institutional ethical procedures and secondly in developing a strategy for the culturally specific ethics that arise out of the digital environment in which the study was for the most part conducted. The second is in making decisions about how to reflect the identities of participants when attributing their public statements.

---

[16] See https://www.mediawiki.org/wiki/Analytics/Research_and_Data/Showcase#April_2015.

Wikipedia content is governed by Creative Commons Attribution Share-Alike copyright license where contributors allow others to republish selections of the work that they contribute to without permission, as long as they provide attribution to Wikipedia. Technically, then, it is legally permissible to cite statements made by editors to their usernames. Scholars like Helen Nissenbaum (2004), however, argue that a person's privacy can be violated when their statements are removed from the social context from which they originated.

The Association of Internet Researchers urge researchers to adopt a notion of privacy as including 'a consideration of expectations and consensus' (Markham & Buchanan, 2012). Since a participant's Wikipedia username can provide personally identifiable information, it is important to enable participants to decide on the level of disguise associated with their statements on the public forums. I ascribe to Amy Bruckman's view (2002) view that Internet users should be seen more as "amateur artists" than 'human subjects' in that users are often people who are using the Internet to express themselves, acknowledging that their usernames have an accompanying reputation. Anonymising the identity of editors should therefore not necessarily be the default approach (and, indeed, anonymising Wikipedians' public discussions would be technically very difficult).

For this reason, I have attempted to share my initial research results with users who I have quoted in their conversations on the wiki platform and on mailing lists, and I have also shared relevant chapters with the people whom I interviewed. I have asked in the release forms how interviewees would like to be named in the study, with all respondents indicating that they would like to be named according to their usernames. All interviewees have been provided with information about the study, along with a

release form where they are asked how they would like to be named in the study (see appendix B).

3.8   Conclusion

This thesis has been designed to produce a rich interpretation of social action on Wikipedia using the methods and principles of ethnography, particularly by scholars who have translated the principles of ethnography in the context of studying online practice. The chapters that follow serve to describe and analyse the features of Wikipedia's socio-technical structure that determines what knowledge is represented on the platform, as well as why some facts travel further than others. The next chapter, in particular, provides an historical analysis of Wikipedia's verifiability principle as the key lens through which encyclopaedic knowledge is filtered. This chapter provides the normative, epistemological framework for understanding daily practice on Wikipedia, a framework that is determined by both social and technological factors.

Chapter 4: Encyclopaedic identity and the verifiability principle

> The threshold for inclusion in Wikipedia is verifiability. The threshold for inclusion in Wikipedia is not truth. (Wikipedia:Verifiability, not truth)

Wikipedia identifies itself first and foremost as an 'Internet encyclopaedia' (en.wikipedia.org/Wikipedia). The two components of this phrase have guided the thinking, strategizing, planning and policy-making for the project over the past fourteen years and constitute the terrain in which facts travel in Wikipedia. There are many sources of uncertainty and risk for users of information in online environments and individuals face uncertainty about the accuracy and credibility of sources of information shared online (Cheshire, Antin, & Churchill, 2010). Features of the online environment, particularly in relation to anonymity, have led to increased perceptions of distrust and uncertainty (Kollock, 1999). It is within such environments of uncertainty that ideologies become important and take hold. According to Geertz (1973), 'It is in country unfamiliar emotionally or topographically that one needs poems and road maps' (p. 218). Road maps, in this non-literal sense, constitute ideological systems that help actors deal with the strain that results from an absence of cultural resources with which to deal with uncertainty.

In the context of Wikipedia, the terrain in which facts travel is constituted by road maps of what constitute encyclopaedias that also determine how they should be designed and constructed. These road maps serve as evidence for editors attempting to re-establish traditional notions of authority and by the editors who wish for Wikipedia to chart new territory for encyclopaedias in the future. For those who want to re-establish tradition,

road maps are provided as evidence of what has worked in the past, and what readers expect from the encyclopaedia. For those who rail against encyclopaedic tradition, road maps of the past are evidence of how the encyclopaedia has failed, and how Wikipedia needs to continue to innovate in order to sustain its mission. In order to understand this tension, we first need to look at Wikipedia's foundational principles, how Wikipedia departs from encyclopaedias of the past and how it exemplifies earlier instantiations.

This chapter discusses Wikipedia's foundational policies, values and principles using a comparison of Wikipedia to other encyclopaedias to highlight a key difference in Wikipedia's conceptualisation and representation of the authors of encyclopaedic facts. Wikipedia's conceptualisation of authorship is centred around the principle and policy of verifiability, which sets up a series of rights claims and communicative acts between readers and editors that I argue are fundamental to understanding decisions that are then made at the edges of the encyclopaedia (at the article level). In order to understand how decisions are made in everyday practice on Wikipedia, it is first important to understand the policy strains editors are faced with. Verifiability, I argue, acts as an ideological system on Wikipedia that has a significant impact on the roles of editors, sources and readers, which subsequently affects what knowledge can be added on Wikipedia.

4.1   Wikipedia's encyclopaedic heritage

Wikipedia aims at representing 'the sum of all human knowledge' and distributing it 'freely to every person on the planet' (Wikimedia Foundation, n.d.). This goal for universality of scope represents a continuation of, rather than a departure from, encyclopaedias of the past. Like Wikipedia, encyclopaedias of the past have aimed at representing all that is known about a certain branch or all branches of knowledge; like

Wikipedia, encyclopaedias of the past have had similar global educational goals. Directed towards the education of the layperson or learner, encyclopaedias have attempted the wide dissemination of what they regard as *useful* knowledge (Burke, 2012; Reagle, 2010).

The term, 'encyclopaedia' comes from the Greek term *enkuklopaideia* for *enkuklios paideia* to mean 'all-round education'. Encyclopaedias, whether general or relating to a specific field of knowledge, have always aimed at comprehensiveness. Comprehensiveness includes the scope of coverage (what knowledge is covered), the depth of coverage (how a subject is covered) as well as how up to date the coverage is (coverage of recent developments) with many encyclopaedias aiming to cover *everything* that is currently known about a particular subject or subjects.

In describing the history and development of encyclopaedias, Collison determines that what is essentially encyclopaedic about a publication is that it aims to reflect the universe of all available knowledge about a particular subject.

> In this article the word *encyclopaedia* has been taken to include not only the great general encyclopaedias of the past and the present but all types of works that claim to provide in an orderly arrangement the essence of "all that is known" on a subject or a group of subjects. This includes dictionaries of philosophy and of American history as well as volumes such as *The World Almanac and Book of Facts*, which is really a kind of encyclopaedia of current information. (Collison, Encyclopaedia Britannica Online Edition, n.d.)

Wikipedia is not the first encyclopaedia with the goal of representing all human knowledge, nor of making that knowledge universally accessible. Joseph Reagle (2010) argues that Wikipedia is heir to the twentieth-century vision of universal access and goodwill of the encyclopaedia advocated by H.G. Wells and Paul Otlet at the turn of the century (Reagle, 2010). Otlet's idea 'foreshadowed Wikipedia' in that it was 'international, multilingual, collaborative, and predicated on technological possibility'

(Reagle, 2010, 24). H.G. Wells was also passionate about the vision of a universal reference work. He hoped that a world encyclopaedia could 'solve the problem of that jig-saw puzzle and bring all the scattered and ineffective mental wealth of our world into something like a common understanding' (Wells, 1938, 8).

The public goal of the encyclopaedia as a tool for general education has been a major feature of encyclopaedic development in the Western world. The Penny Cyclopaedia published between 1833 and 1843, for example, was authored by the members of 'The Society for the Diffusion of Useful Knowledge'. The encyclopaedia aimed to:

> give such general views of all great branches of knowledge, as may help to the formation of just ideas on their extent and relative importance, and to point out the best sources of complete information. (Society for the Diffusion of Useful Knowledge, 1833)

Along with comprehensiveness, encyclopaedias have asserted general educational goals by employing vocabulary aimed at the layperson rather than the expert. Encyclopaedias are designed to enable a particular type of information seeking, for use as a reference among readers who are looking up a particular subject, rather than being read from start to finish. Paper encyclopaedias are generally arranged in alphabetical order, but digital encyclopaedias employ search functionality in order to more efficiently enable this kind of targeted information-seeking practice. Paper encyclopaedias' 'see also' items at the end of entries and digital encyclopaedias' hyperlinks to related entries, for example, offer opportunities for cross-referencing.

Common to the envisioning of the encyclopaedia as the sum of all human knowledge is a vision of the layperson, the non-expert, or the learner, as the receiver of knowledge. Although encyclopaedias have, in the past, purported to represent *all* knowledge, they have never sought to represent *everyone's* knowledge. The fact that there is a single

entry to represent each phenomenon in the encyclopaedic corpus rather than a series of alternative descriptions written by multiple authors is reflective of the necessity of a dominant perspective and the affordances of paper publishing.

Historically, the characteristics of encyclopaedias have changed as social norms dictating what is considered useful knowledge, and how knowledge can be arranged and retrieved using new technologies, have changed. As a consequence, what is technically possible may drive what content is regarded as essential to the encyclopaedia, and what is socially normative may drive what the technology is tasked with achieving. Such visions of what is essential to the encyclopaedia and what should not be included have imposed limits on the types of knowledge encyclopaedias have included in their coverage.

According to Collison, what has been considered encyclopaedic has varied throughout history.

> The degree of coverage of knowledge has varied according to the time and country of publication. Illustrations, atlases, and bibliographies have been omitted from many encyclopaedias, and for a long time it was not thought fitting to include biographies of living persons. Indexes are a late addition, and most of the early ones were useless. Alphabetical arrangement was as strongly opposed as the use of any language but Latin, at least in the first 1,000 years of publication in the West, and skilled group editorship has a history of some 200 years. (Collison, 2015)

Collison's statement indicates that different kinds of content and style have been considered appropriate for encyclopaedias over time. Such beliefs have necessarily imposed a limit on what types of knowledge can be represented. Another limit to what has been represented by encyclopaedias lies in the attitudes of the encyclopaedic producers towards the knowledge of different people. Burke (2012), for example, notes that the dissemination of knowledge to all was encouraged and celebrated from the 18th century onwards, but this dissemination went from educated, Western, male knowledge

holders to the poor, to women and to people living in the colonies. The spread of Western knowledge was encouraged as part of the imperial enterprise because Westerners believed their knowledge to be superior (Burke, 2012, p. 212). According to Thomas Macaulay's 'Minute on Indian education', for example, 'a single shelf of a good European library was worth the whole native literature of India and Arabia' (Thomas Macaulay, 1831).

In the same way that the content of the Penny Cyclopedia reflected attitudes regarding who were perceived as the experts and subjects of education in nineteenth century England, so too does Wikipedia's content (and the limitations to that content) reflect attitudes and conventions about who are perceived as the experts and subjects of knowledge today. Where Wikipedia departs significantly from encyclopaedias of the past, however, lies in its conceptualisation of authorship. Whereas encyclopaedias of the past generally tended to acknowledge the authors of individual entries but not to provide evidence for every statement contained within them, the authors of Wikipedia articles are unacknowledged in the article while individual statements tend to be heavily cited to external sources. In the past, encyclopaedias were signed as a way to lend credibility to the information if they were written by experts in a relevant subject. In other cases, authors avoided signature because they feared opposition or ridicule (Loveland & Reagle, 2013). Wikipedia lies in stark contrast to past encyclopaedias in its profuse (and growing) use of in-text citations and sources for further reading, especially on larger and more mature versions such as English and Polish Wikipedias (Jemielniak, 2014).

Encyclopaedia Britannica's online entry on *encyclopaedia*, for example, consists of 24 pages organized across four main sections written by two expert authors. The section

titled 'Encyclopaedias in general' in this example is written by Robert L. Collison who is named as the author at the top of the article (see figure 4.1 below). Collison's name is linked to his profile page where his credentials are provided as:

> Professor of Library Science and Information Studies, University of California, Los Angeles, 1968–73. Author of *Encyclopaedias: Their History Throughout the Ages* and others. (Encyclopaedia Britannica, n.d.)

The format and content of Encyclopaedia Britannica's attribution – which lists Collison's professorial credentials, his institutional position and his authorship of a work of scholarship in the category of the subject of the article – serve to authorize Collison to speak about the subject of encyclopaedias by virtue of his noted expertise.

Figure 4.1 Screenshot of the entry for 'encyclopaedia' on Encyclopaedia Britannica online

Source: http://www.britannica.com/topic/encyclopaedia (2 July 2015)

Wikipedia, on the other hand, features no named authors in the main text of the article titled *encyclopedia* (see Figure 4.2). The authors of the article can only be found by clicking through each page in the history tab in order to see the username associated with each edit over time. Instead of listing authors in the text of the article or even providing a list of authors hyperlinked from the article, Wikipedia's focus regarding attribution is on providing multiple citations to the sources of statements within the main text of the article. The sentence below, for example, contains not one but four in-text citations:

> Generally speaking, unlike dictionary entries, which focus on linguistic information about words, encyclopedia articles focus on factual information to cover the thing or concept for which the article name stands. [4][5][6][7] (en.wikipedia.org/encyclopedia)
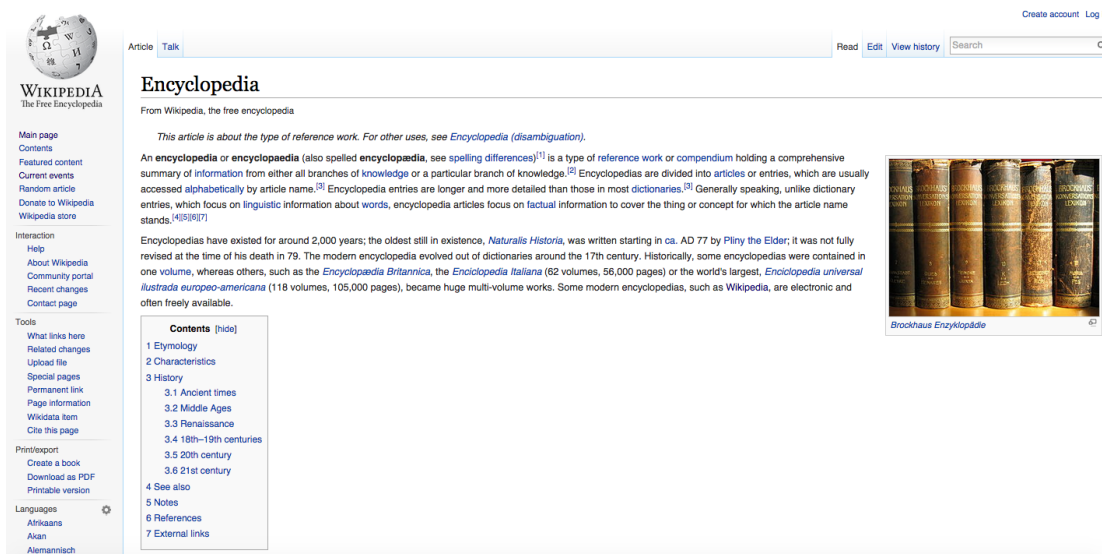
Figure 4.2 Screenshot of the Wikipedia entry for 'encyclopaedia'

Source: https://en.wikipedia.org/w/index.php?title=Encyclopedia&oldid=675180187

It is notable that the Britannica entry only contains citations for direct quotations from primary sources and does not provide URLs or other location information in order to assist the reader in looking up those sources. The sentence below, for example, contains

an evaluative claim that would almost certainly either be removed as 'unsourced' or appended by the superscript text: 'citation needed' if it were to appear in Wikipedia.

> Even today a modern encyclopaedia may still be called a dictionary, but no good dictionary has ever been called an encyclopaedia. (Collison, 2015)

The difference in ways in which Wikipedia and Britannica attribute sources and authors is not only a stylistic issue. This difference represents a fundamental shift in the ways in which Wikipedia derives its authority and consequently how it defines the relationship and identities of its primary stakeholders including sources, editors and readers. The conventions relating to authorship and attribution on Wikipedia are dictated by policies that have been developed to specify not only stylistic conventions but also epistemological principles. A central principle which finds its way into multiple policy documents, norms, technical tools and stylistic elements on Wikipedia is the principle of verifiability.

4.2   Verifiability

In Wikipedia's early years, contributors developed a series of policies that would guide what Wikipedia would include or exclude. Starting with the development of the Neutral Point of View (NPOV) policy in the first few months of Wikipedia's founding in 2001, numerous policies were added in order to clarify what was acceptable on Wikipedia, with each policy increasing extensively in length and complexity (Butler, Joyce, & Pike, 2008).

Wikipedia's current policy environment is now constituted by hundreds of documents including policies, guidelines and essays that enjoy varying levels of authority among editors. Policies are considered the highest authority on Wikipedia and violations are considered grounds for user sanction including banning or disciplinary action by

administrators. Guidelines are considered to be less official than policies and have numerous exceptions. Essays are written by individual editors or small groups of editors and can become guidelines or even policy if they obtain enough support.
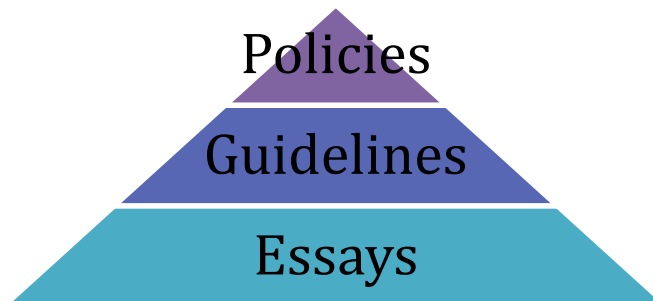
Figure 4.3 Hierarchy of rules on Wikipedia

Source: adapted from Wikipedia:List_of_policies

Today there are 56 English Wikipedia policies, about a hundred guidelines and hundreds of essays. Policies cover issues relating to the scope of the encyclopaedia and what is considered suitable for inclusion, as well as the conduct of editors, deletion issues, enforcement issues, legal considerations and procedural issues (Wikipedia:List_of_policies). All language editions must comply with Wikipedia's core policies but there are differences in the implementation of policies relating to copyright law, for example.[17]

Wikipedia's content policies are centred around three core principles: neutral point of view (NPOV), no original research (NOR) and verifiability (see figure 4.4). NPOV demands that articles should be written without bias by fairly and proportionately representing all significant views. The 'no original research' policy requires that

---

[17] Different national interpretations of exceptions and limitations to copyright mean that rights to produce material for educational purposes on Wikipedia are wider in some countries than others.

Wikipedia editors do not publish original thought and that all material must be attributable to a reliable source. Verifiability determines that all material challenged or likely to be challenged must be attributed to a reliable source (Wikipedia:Core content policies).

Core content policies
Neutral point of view
No original research
Verifiability
Other content policies
Article titles
Biographies of living persons
Image use policy
What Wikipedia is not

V · T · E

Figure 4.4 Core content policies from English Wikipedia

Source: Wikipedia:Core content policies

The profuse use of in-text citations on Wikipedia is reflective of strains caused by Wikipedia's verifiability policy. According to the verifiability policy, 'all quotations and any material challenged or likely to be challenged must be attributed to a reliable, published source using an inline citation' (Wikipedia:Verifiability). The default approach, according to policy authors, should be to provide a citation for material and that material that is un-sourced may be removed by this criterion alone.

> All material in Wikipedia mainspace, including everything in articles, lists and captions, must be verifiable. All quotations, and any material whose verifiability has been challenged or is likely to be challenged, must include an inline citation that directly supports the material. Any material that needs a source but does not have one may be removed. (Wikipedia:Verifiability)

In addition, content that is especially controversial requires 'multiple high-quality sources' (Wikipedia:Verifiability). Reasons to include multiple sources are wide-ranging, according to policy, and are characterized by claims that are 'surprising or apparently

129

important' but are 'not covered by multiple mainstream sources' or else claims that are 'challenged' and 'supported purely by primary or self-published sources or those with an apparent conflict of interest'. The requirement for multiple sources also applies to 'reports of a statement that seems out of character, or against an interest they had previously defended' as well as to

> 'claims that are contradicted by the prevailing view within the relevant community, or that would significantly alter mainstream assumptions, especially in science, medicine, history, politics, and biographies of living people. This is especially true when proponents say there is a conspiracy to silence them. (WP:Verifiability)

The wide scope of these 'red flags' (WP:Verifiability) requiring multiple, high quality sources means that almost all the statements that occur on Wikipedia would be subject to their application.

Wikipedia's verifiability policy can be seen as a merging of the traditions and epistemologies of both the scientific field and the free and open source software movements, both of which have had a significant impact on Wikipedia's norms and practices. Whereas science is developed according to the original contributions of its practitioners, Wikipedia is authored by contributors who are banned from doing original research. Whereas software can be written in different ways, those differences do not have a political impact in the same way that encyclopaedic facts do. These and other differences between Wikipedia and science and Wikipedia and software have resulted in a unique implementation, function and implications of citation practice for the representation of knowledge on Wikipedia.

Academic institutions teach students norms of referencing and attribution through a system of reward and punishment. Academics and scientists are held to these rules by facing sanctions by the academic community if they are not followed. The process of

citation and referencing in academia serves practically to distinguish the ideas of the author from those being cited, and is based on the principle that science advances according to iterative steps. The metaphor of dwarfs standing on the shoulders of giants is based on the principle of discovering truth by building on previous discoveries, an aphorism commonly attributed to Sir Isaac Newton with his phrase:

> If I have seen further, it is by standing on the shoulders of giants. (Newton, 1959)

Citation and referencing play a role in maintaining the authority of scientists by establishing the individual author's contribution to the production of knowledge, thus placing authority in the individual scientist as separate and distinct from representations by others. Copyright laws and plagiarism norms protect the right of the individual author and reflect society's focus on the individual rather than the collective in terms of authorship.

Whereas the goal of science is to advance knowledge, the goal of Wikipedia (and encyclopaedias in general) is to make knowledge accessible. Unlike science, the emphasis of citation policy for Wikipedia is on distinguishing the author's perspective from other sources, not in order to enable evaluation of the unique component but in order to remove the unique component. Original research is banned on Wikipedia and editors are allowed to remove material that is unsourced. Like it does for science, Wikipedia's citation practice provides a mechanism for establishing its authority, but in Wikipedia's case authority originates with the named source rather than the hidden author of the article.

The function of citation policy on Wikipedia is as a practical mechanism for ensuring copyright compliance and for establishing and ensuring the rights of the reader to verify

statements. Verifiability is framed as a way of empowering readers, making transparent the path by which information came to Wikipedia and providing the ability for readers to check whether claims on Wikipedia are backed up by reliable sources. The verifiability policy's orientation towards the reader is emphasized in the way that the three core content policies are written. Whereas the NPOV policy tells editors how they should write about a particular subject and the NOR policy tells editors what they should exclude, the verifiability policy is framed in terms of the right of the reader to be able to check the source of information found on Wikipedia.

> Readers must be able to check that Wikipedia articles are not just made up. This means that all quotations and any material challenged or likely to be challenged must be attributed to a reliable, published source using an inline citation. (Wikipedia:Verifiability)

The act of the reader checking that the sources are being accurately reflected in the text of the article is articulated as politically important to Wikipedia and is a significant feature of Wikipedia's claim to authority. In Isin and Ruppert's (2015) terminology, the verifiability establishes a rights claim which establishes that readers have a right to verify information on the page.

Wikipedia's verifiability policy is implemented according to a series of practices, norms and tools. One key practice is the ability of readers to add 'citation needed' templates for content that they are sceptical of. Adding the 'citation needed' tag to the end of a dubious sentence on Wikipedia is a signal to the editor who added the material to prove her/his claims using citations to reliable sources. According to policy (shortened as Wikipedia:ProveIt), editors contributing material are responsible for providing evidence to support that material.

> The burden to demonstrate verifiability lies with the editor who adds or restores material, and is satisfied by providing a citation to a reliable source that directly supports the contribution. (Wikipedia:Verifiability)

The process of requesting further evidence and providing evidence constitutes a communicative act within Wikipedia's socio-technical environment, a type of call and response that is enacted in the name of the readers for whom editors are working. Although the verifiability policy is framed in terms of the rights of readers to ask for further evidence, in practice it is other editors who claim this right, often on behalf of readers. This is because, in order to engage in the communicative action of asking for further evidence, readers must know how to produce the *citation needed* trace. Along with a number of additional traces that embed the norms, values and identities of Wikipedians in their expression, the 'citation needed' tag enables a series of communicative actions by those who deploy it.

In order to request further evidence, a reader would need to know, not only that they have a right to request further evidence, but also how to produce the trace by applying the template, as well as the norms dictating whether such a trace is required or not required in the context of the article. User documentation about the template for the *citation needed* tag (Template:Citation_needed) provides the syntax needed to produce the trace:

> {{citation needed|reason=*Your explanation here*|date=July 2015}}

The syntax is followed by a paragraph about "When not to use the template" including occasions when the content refers to living persons and should be removed immediately rather than being queried, when more specific tags should be used (a list of 40 other tags are provided), when the editor should rather find the source themselves, and when sources are not required, that is, for 'common, well-known facts', which links to a Wikipedia article on 'common knowledge' (en.wikipedia.org/Common_knowledge).

This articulation of rights and communicative actions is a foundational feature of the free and open source software movement in which the freedom to reverse engineer the software has been seen as one of the core freedoms for software users. Four freedoms are outlined by the free software 'General Public License' (GPL) agreement including the freedom to run the program, to study and change the program, to redistribute copies and to improve the program and release improvements to the public.[18] The principle of verifiability is related to all four freedoms which also constitute obligations because of their contractual format. In order for a reader to be able to check whether information is accurately reflected on Wikipedia, they need to be able to access the information, check whether the source is accurately reflected in the text of the article and then change the Wikipedia article if they believe that it needs to be improved.

The free software movement has had a significant influence in the development of Wikipedia. Wikipedia founder, Jimmy Wales, has recognised this influence on his leadership of the project, noting that he was particularly interested in Eric Raymond's text about free and open source software, 'The Cathedral and the Bazaar' (Schiff, 2006). Wikipedia's first copyright license was the GNU-FDL (free documentation license) written by the Free Software Foundation originally for software documentation.

There have always been vibrant and intense debates about copyright licensing within the Wikipedia community. As a result of Wikipedia's move from the GNU-FDL to Creative Commons in 2009, attribution requirements for re-publishers were lightened. The FDL required the attribution of every author of a text including the contributions of

---

[18] See http://www.gnu.org/licenses/gpl-3.0.en.html.

that author to the text when the work was copied, re-published or attributed. According to the Wikimedia Foundation, this would have effectively required complete duplication of information in the history view of the article which proved significantly onerous (Wikimedia Foundation, 2009). There was strong debate within the community about whether to make the change or not, but eventually the Wikimedia community approved the migration of licenses with a community vote in 2009 (Wikimedia Foundation, 2009).

This meant that no longer would attribution require the entire list of authors of Wikipedia articles and their contributions but rather that attribution could take place in the manner specified by the license holder. Wikimedia editors ceded their copyrights to the Wikimedia Foundation in order to facilitate the change.

The rights claim that is represented by Wikipedia's attribution conventions, particularly around the use of the "[citation needed]" tag, is so significant that it has spilled out into the realm of popular culture. Popular Internet cartoonist Randal Munroe, for example, published a cartoon in 2007 that featured the "[citation needed]" tag (see figure 4.5).
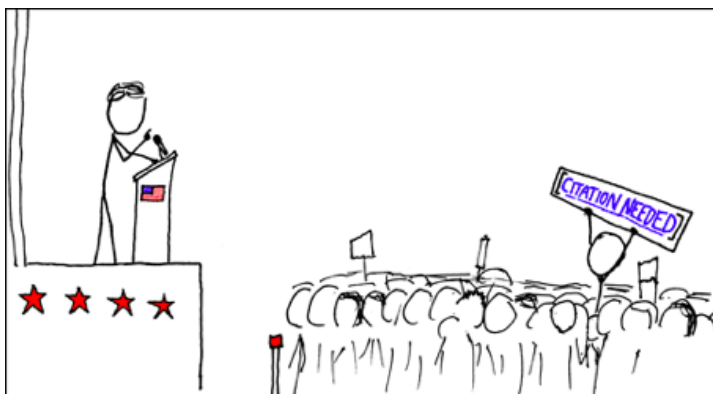

Figure 4.5 'Wikipedian protester' by Randal Munroe, xkcd

Source: http://xkcd.com/285

The cartoon shows a member of the public responding with scepticism to a politician making a speech on a podium.

135

> Cueball holds up a sign showing CITATION NEEDED during a political speech. The sign text is a Wikipedia template placed next to a statement that needs a citation, usually because of questionable validity. Cueball is obviously using this template to challenge the politician's speech. (Explain xkcd, 2015)

Munroe succeeded in linking the liberatory action enabled by the deployment of the phrase to the context of mainstream. The cartoon became so popular that it was translated into 16 different languages, [citation needed] stickers were placed on billboards in New York by Matt Mechtley in order to question dubious claims being made by advertisers and the phrase has been used on posters during political protests.[19] Mechtley writes that he started the project to place stickers with [citation needed] on billboards in New York City as a way of questioning dubious claims.

> One of my favorite quirks about [Wikipedia] are the little [citation needed] tags that users can place in an article, indicating that a dubious claim needs a reference. One day an idea struck – what statements are more dubious or outright ridiculous than those in advertisements? Thus, an OM project was born. I had 250 8×2 inch stickers printed, which I handed out to friends, who circulated them further. In true wiki fashion, the final placement of the stickers is a collaborative effort, now distributed and anonymous. If anyone sees one somewhere, please make a photo! I've been tagging my photoset on Flickr with citationneeded and wikiffiti. (Mechtley, 2008)

While Munroe used the [citation needed] statement to speak back to what he believed was missing in mainstream politics, Mechtley was using the [citation needed] statement as a way of speaking back to the practices and conventions of corporate advertising which he regarded as similarly dubious. The phrase provided Mechtley with the language in which to express his scepticism and to do it in a way that would resonate with others who knew the meaning of the phrase. Emerging originally from the free and open source software field as a politically charged remix of traditional academic norms, Wikipedia's [citation needed] phrase has spilled out into the world outside of Wikipedia (see figure 4.6).

---

[19] See http://knowyourmeme.com/memes/citation-needed.

The verifiability principle from which the phrase is derived is not only a rule that applies to the way that statements are structured on Wikipedia; it is an ideological system that is constituted by a series of values, principles, policies, and norms. Verifiability effectively defines the communicative acts that are possible within Wikipedia's socio-technical system.
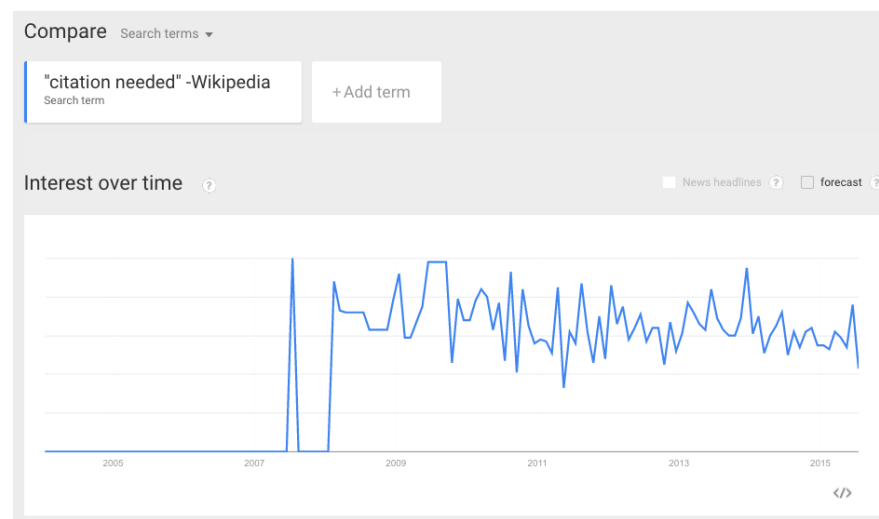
Figure 4.6 Google Trends report showing the use of the term 'citation needed' (outside of Wikipedia)

Source: https://www.google.com/trends/

4.3   Verifiability as ideology

According to Clifford Geertz, ideologies are cultural symbol-systems that serve as extrinsic sources of information in situations where institutionalized guidelines for behaviour are lacking (Geertz, 1973, p. 218). In the context of Wikipedia, gaps are created by the decentralised environment where there is no central editorial oversight, and the online environment where the boundaries between knowledge producing fields such as science and journalism are themselves unclear and in the process of revision (Anderson, 2013; Lewis, 2012). The principle of verifiability has been constructed in order to fill in these gaps. Verifiability, in Geertz's terms, acts as a 'matri(x) for the

creation of collective conscience' (p. 220) by providing a framework by which people can collectively recognise true knowledge and expel false knowledge.

The concept of verifiability has its origins in the development of the scientific profession. Stephen Shapin (1984), for example, discusses the way in which Robert Boyle's public experiments in pneumatics using the air pump in 1650s and 1660s England were legitimated. Numerous controversies resulted from Boyle's first experiments because there was initially a lack of consensus about what constituted scientific knowledge.

> What was of issue in the controversies over Boyle's air-pump experiments during the 1660s was the question of how claims were to be authenticated as knowledge. What was to count as knowledge, or 'science'? How was this to be distinguished from other epistemological categories, such as 'belief' and 'opinion'? What degree of certainty could be expected of various intellectual enterprises and items of knowledge? And how could the appropriate grades of assurance and certainty be secured? (Shapin, 1984, p. 482)

According to Shapin, Boyle employed three technologies in order to institutionalise knowledge, to 'make common' (p. 482) the methods by which knowledge could be authenticated. These included the material technology of the air-pump that produced the experiments, a literary technology whereby the phenomena produced by the pump were made known to those who weren't direct witnesses, and the social technology of conventions people should employ in dealing with one another and considering knowledge claims.

Boyle's experiments represent 'a revolutionary moment in the career of scientific knowledge', according to Shapin (p. 482). The choice of authoritative witnesses was not, however, revolutionary. Boyle called for witnesses to the air-pump experiments from the upper classes (selecting Oxford professors rather than Oxfordshire peasants, for

example). What was revolutionary was the construction of the experiment and its inscription by particular witnesses as authoritative. Witnesses then wrote about their experience in a realistic style which had the effect of 'multiplying authority by multiplying witnesses' (Shapin, 1984, p. 488). The air-pump functioned as an objectifying resource.

> The machine constitutes a resource that may be used to factor out human agency in the intellectual product: "it is not I who says this: it is the machine that speaks," or "it is not your fault; it is the machine's". (Shapin, 1984, p. 508)

Shapin's focus on the importance not only of producing knowledge but of communicating knowledge is useful to highlight that methods for the public verification of knowledge have been necessary to the generation and validation of facts. Such methods (consisting predominantly of literary techniques) needed to be constructed in a manner conducive to the distribution of authority in the 17th century England. Authority for witnessing was designated to the landed gentry, but the machine itself was also deemed to be authoritative. The partial loss of agency as the responsibility for establishing truth claims in the case of Boyle's air-pump was effectively being handed over to the machine. Despite this, the role of human witnesses was equally significant in inscribing the actions of the machine for the public constitution of knowledge.

The process of distributing authority has continued, albeit in different arrangements, today. Wikipedia's policy documents frame knowledge in a way that distributes responsibility for the truth-making process among readers, sources and editors. In the same way that the machine in Boyle's time served as a means of 'shifting agency onto natural reality' and thereby making it appear to be an independent fact of nature, Wikipedia shifts agency onto natural reality by virtue of its verifiability claims that are then implemented or inscribed using machines.

The verifiability of a phenomenon is the extent to which it can be proved to be true, and the act of verification is constituted by the action of demonstrating or proving that something is true by means of evidence. Wikipedia's framing of verification is a very specific instantiation of this principle. Verifiability as framed by Wikipedia policy is about enabling the reader to check whether the source has been accurately reflected in the article, rather than whether the source is reliable or whether it is accurate according to her/his own observation of a phenomenon. Verification within Wikipedia's socio-technical framework involves a process of proving not that a claim is *true*, but that a claim is a truthful representation of an independent source. This instantiation of verification is the opposite of Boyle's experiments whereby the inscriber of knowledge was also the person who experienced the phenomenon and where others believed them because of their ability to render their experience palpable to readers.

In an article about the game of cricket, for example, a statement on Wikipedia might read that 'the game of cricket is played between two teams of 11 players each' with a citation to the website of the International Cricket Council. The statement is verifiable in Wikipedia's terms because readers can check for themselves whether the source has been accurately represented. In other words, verifiability is enabled because the reader can check whether the source is accurately summarised in the article by looking up the source, reading it and then comparing her/his understanding of the meaning of the source to the way it is represented in the article. Verification is not about the ability of the reader to determine whether, according to her/his own observations, the rules of the game of cricket have been accurately reflected in the article, or whether the website of the International Cricket Council constitutes reliable evidence for the statement being

cited. It is simply about the ability of the reader to check whether the source is being accurately represented in the article.

An alternative perspective (indeed one that is held by one of the stakeholders in the debates to follow) is that verifiability should be concerned with equipping readers with the tools necessary to verify whether the claim made in the text is, in fact, true according to their own observations.

In the same way that Boyle's air pump enabled early scientists to absolve themselves of responsibility for determining the truth by pointing to the independence of the machine, Wikipedia's verifiability principle is framed in a way that means that Wikipedia editors don't have to take responsibility for the truth of their claims because they assert that the claims aren't asserted by the editors themselves, and are rather created by the sources that editors are merely passively reflecting (and attributing) in the text of the article. In fact, Wikipedia explicitly states in its policy that the primary criterion for the inclusion of information in Wikipedia is verifiability, not truth (Wikipedia:Verifiability). There is effectively an attempt to evade questions of truth by asserting that responsibility is, instead, in the hands of reliable, independent sources.

The verifiability principle sets up particular roles for editors, sources, readers and subjects on Wikipedia. Editors are banned from doing original research and must cite every claim made in an article that has been or could be challenged with a relevant source. The implication of policy is that Wikipedia editors are merely passive aggregators of information already published by external sources. The authors of sources and citations, on the other hand, must be located outside the sphere of the encyclopaedia; otherwise these authors will be regarded as having a conflict of interest.

The source's independence from the phenomenon and from Wikipedia is an important criterion for determining their reliability and thus their acceptability for Wikipedia.

Furthermore, readers are protected from potentially inaccurate information by the availability of citations that enable readers to request that a source be provided and then check whether information from the source is accurately reflected on Wikipedia. Similarly, the subject of the article needs to also remain outside Wikipedia's sphere since any involvement by the subject constitutes a potential conflict of interest. Instead of a free for all, Wikipedia policy already sets up clear boundaries between who can participate in Wikipedia's representation of knowledge.

Editors are framed as passive reflectors of existing knowledge rather than creators of knowledge themselves. Editors, the authors of Wikipedia's articles, take a backstage role in the presentation of phenomena. Instead of being named as part of the article as in the case of Britannica, authors are hidden behind the page and there is no way of obtaining a list of contributors to each page without clicking through the history pages. The term *editor* itself (rather than *author)* as the predominant terminology for article contributors is significant because it highlights the perceived role of editors as passive aggregators, rather than active creators, of information.

Furthermore, Wikipedia's verifiability policy defines sources as either the type of work, the creator of the work or the publisher of the work (Wikipedia:Verifiability), thereby omitting the article's editors as sources of information contained within the article. According to the verifiability policy, contributors don't actively create a work of scholarship but merely select and prepare information already created. Aggregation, in other words, is represented as a technical process rather than an interpretive skill.

In the same way that Boyle framed the air pumps' activities as an objective demonstration of knowledge, Wikipedia's verifiability ideology is the key means by which Wikipedia derives its authority. Wikipedia's authority to represent a particular subject originates from the sources that it uses to construct the article, rather than from the authors of the article itself. While Encyclopaedia Britannica derives its authority from the credentials of the experts that it employs to write articles about phenomena in their areas of expertise, Wikipedia asserts that the authors of articles need not be subject matter experts because the information is merely being collated from the published work of experts.

The only way that Wikipedians can claim authority to represent quality information without the help of experts is by downplaying their own creative role in the process of knowledge construction. At its heart, this practice reflects a deep insecurity in the ability of the crowd to be trusted with representing knowledge. In Jasanoff's terms, Wikipedia's verifiability principle is a reflection of how traditional expertise is reaffirmed, while at the same time establishing the basis for new sources of expertise as they become recognised. Although the scientific discourse initiated by Robert Boyle was led by the reaffirmation of the authority of the landed gentry, over time those who inscribed the workings of the air pump and its successors became authorities in themselves. Similarly, we are witnessing how on Wikipedia, at the same time that editors use sources as claims to legitimacy, they are becoming authorities and experts themselves.

The focus of policy on adding only information that can be verified by an independent, reliable source has led to a series of unintended consequences. Firstly, information that is easily verifiable tends to be added to Wikipedia so that there are situations where

knowledge that isn't verifiable or at least easily verifiable remain outside the corpus. Secondly, the focus is on adding sources as trust signals rather than seeing to it that sources are maintained over time so that the number of dead links contained in citations is highly significant. Both these effects indicate that the promise of verifiability as a means for public engagement in establishing the veracity of truth claims can become empty at best and a tool for manipulating information according to the interests of the editor at worst.

Many of the more mature Wikipedias such as English, German and Polish Wikipedias have, over the years, increased their strictness regarding sources (Jemielniak, 2014). According to Jemielniak, 'In some cases the results are absurd' (p. 21). Jemielniak recalls a situation in the Polish Wikipedia in which all female members of parliament were described using the male forms of words. He wanted to correct it but an administrator argued that the Polish parliamentary acts used only male nouns and the acts were the authoritative sources.

> I argued that such acts are always written with just one form for ease of reading. However, my argument was moot without valid published sources using female forms of the positions, and I was not able to find any. Thus, I decided to create a source: through my publisher I contacted a well-known professor of linguistics and asked him to write an opinion in an online language advisory portal he ran, under the publisher's auspices. He agreed, and the opinion, recommending the use of female forms for female MPs, was published; no one could now object to changing the inboxes, and they were immediately amended. (Jemielniak, 2014, p. 21)

Jemielniak notes that, in order to respond to the absurdity of situations like this, editors often have to take strategic steps in order to produce the evidence for their claims through others. This provides the appearance of independence when actually this evidence has effectively been produced at least partly by the editor themselves. For those who aren't aware of what the ideal source is on Wikipedia and what requires

sourcing, getting someone appropriate to write about challenged content is too high a barrier.

On Wikipedia 'Even if you're sure something is true, it must be verifiable before you can add it.' (Wikipedia:Verifiability) This principle is often summarized as 'Verifiability, not truth' as the threshold for inclusion of content in Wikipedia. The essay by the same title (Wikipedia:Verifiability, not truth) is linked to the verifiability policy and explains the principle noted in the opening of this chapter:

> The threshold for inclusion in Wikipedia is verifiability. The threshold for inclusion in Wikipedia is not truth. (Wikipedia:Verifiability, not truth)

Critics of the policy have argued that it constitutes a "license to publish lies provided someone else did it first" (Hersch, 2012). An essay created by [Kotniski] called 'Truth, Not Verifiability' (Wikipedia:Truth,_not_verifiability) attempts to correct what s/he sees as undue emphasis on verifiability and that the consequences of the policy's wording are that editors only add information that is verifiable rather than that which is true.

> The *de facto* primary criterion for the inclusion of information in Wikipedia is **truth, not verifiability** – whether reliable sources state it to be true, not whether individual editors think they can verify it.
>
> You may have noticed that the *de jure* primary criterion, as stated on Wikipedia:Verifiability, puts this the other way round: "verifiability, not truth" – whether material can be verified by reliable sources, not whether individual editors believe it is true.
>
> It does not really matter. The distinction being made here is not really between truth and verifiability at all, but between the statements made by reliable sources (which we want to include in the encyclopedia), and the unsupported claims of Wikipedia editors (which we don't).
>
> It's not quite as simple as that, of course, but once you know that Wikipedia is supposed to reflect what reliable sources say, not its editors' private thoughts and unsourceable personal knowledge, most of the rest is more or less common sense. (Wikipedia:Truth,_not_verifiability)

Sources are essential objects to Wikipedia's functioning, both because they are necessary for the inclusion of individual facts according to policy, but also because they affect the perceived authority of an article. Research experiments with Wikipedia readers have found that users consider the presence of sources or references to be one of the three most important aspects of the trustworthiness of Wikipedia articles (Lucassen & Schraagen, 2010) but that as long as sources were present, the quality of the sources did not seem to matter (Lucassen & Schraagen, 2013). It is only when users are suspicious of the source of the information and thus perform a thorough, systematic evaluation, that the quality of the sources influences trust. Otherwise, heuristic evaluation is the dominant strategy, even when users are specifically asked to evaluate credibility. This phenomenon has been termed 'reference blindness' (Lucassen & Schraagen, 2013).

An article with numerous sources is therefore more likely to be accepted as credible by readers. In order to try to control the acceptance of particular content, editors can use sources strategically as a signal of authority, rather than as the substance of credibility claims. One of the studies that I conducted with colleagues (Ford et al., 2013) indicated that a significant proportion of Wikipedia's citations are to dead links. Dead links prevent the reader's ability to check up on how information from the source is being reflected, but in many cases, readers don't perform such checks in any case, taking the shortcut by only checking whether citations appear in the article.

Verifiability, then, has the potential of becoming an empty performance where all the right signals appear in the article but there is no substance or depth to the performance. Readers don't tend to look up sources, but even if they did, the process is often stilted

because URLs are dead or unavailable, or else the sources have been strategically chosen in order to advance the views of individual or groups of editors.

4.4 Ideological strains

According to Geertz, ideological strains are endemic to society because 'no social arrangement is or can be completely successful in coping with the functional problems it inevitably faces.' (Geertz, 1973, p. 203) In this context, ideological strains exert themselves in 'insoluble antinomies', discontinuities between norms in different sectors, discrepancies between different goals and in contradictory role expectations (p. 203). Although Wikipedia's verifiability principle has provided a matrix for attempting to solve problems over what constitutes reliable knowledge, it cannot entirely solve these problems because it hinges on editors being able to identify what a reliable source is.

According to Wikipedia policy, great emphasis is placed on the use of reliable sources but the policy provides little practical advice about how to recognise a reliable source beyond features such as a publication's 'history of fact checking' (Wikipedia:Identifying reliable sources). This lacuna exists because the appropriateness of a source or citation is almost completely contextually determined.

Policy states that the word 'source' refers to "the piece of work itself (the article, book), the creator of the work (the writer, journalist), and the publisher of the work (for example, Random House or Cambridge University Press) and that any of the three can affect reliability (Wikipedia:Identifying reliable sources).

> Reliable sources may be published materials with a reliable publication process, authors who are regarded as authoritative in relation to the subject, or both. (Wikipedia:Identifying reliable sources)

The policy does not define what a reliable source is, but instead describes a rough hierarchy of what constitutes reliable sources from 'self-published sources' which includes personal web pages and Tweets which 'should generally not be used', to third-party, published sources with a reputation for fact-checking and accuracy which 'should be used as often as possible'.

For each of these media, the policy highlights exceptions. Self-published sources *can* be used to support information about a person or organization being described in articles, and opinion pieces from third-party, published sources might *not* be used if they are not subject to sufficient editorial oversight. An article about President Barack Obama, for example, might use a Tweet from his verified Twitter account as an example of his unofficial position on a particular issue. Mainstream news reporting is 'generally considered to be reliable for statements of fact, though even the most reputable reporting sometimes contains errors' (Wikipedia: Identifying reliable sources). Sometimes news reports are used to back up facts; sometimes they are treated as primary sources and therefore unreliable.

Secondly, policy indicates that articles need to be based 'largely on secondary sources'. It defines secondary sources as third-party sources that are one step removed (whether by time or location from an event), whereas primary sources are from eyewitness accounts directly involved in an event. According to policy, tertiary sources can be used to support broad summaries of topics especially when secondary sources contradict one another.

Such a division of sources can become complicated in the context of an article. Journalism can be considered a secondary source if it provides in-depth analysis but a

primary source if it provides opinion or eyewitness accounts and Wikipedia can be used as a tertiary source for an article about a topic outside of Wikipedia, but it is a primary source in an article about Wikipedia. There is resistance among many Wikipedians to ever using primary sources, but as the authors of the essay below write, '"Primary" is not another way to spell "bad"':

> Primary sources may only be used on Wikipedia to make straightforward, descriptive statements that any educated person—with access to the source but without specialist knowledge—will be able to verify are directly supported by the source. This person does not have to be able to determine that the material in the article or in the primary source is True™. The goal is only that the person could compare the primary source with the material in the Wikipedia article, and agree that the primary source actually, directly says just what we're saying it does. (Wikipedia:Identifying and using primary and secondary sources)

Although policy talks about different types of media (for example, books and academic journals) being more reliable than others, it continually refers to the fact that 'the reliability of a source depends on context' (Wikipedia: Identifying reliable sources). Wikipedia policy does not constitute hard and fast rules about what should be included or excluded, and there is a great deal of complexity about which rules apply to which types of content (Kriplean et al., 2007). Two areas in which boundary disputes are on-going are the coverage of breaking news events and the coverage of topics for which few (if any) published sources exist.

Although policy determines that 'Wikipedia is not a newspaper' and should not include news reports or journalism, the boundaries of journalism itself are undergoing significant change (Lewis, 2012) and Wikipedia articles concerning breaking news topics are actually the most popular to read and write about on Wikipedia (Keegan, 2013). Debates about whether a particular event should be included on Wikipedia or not are often surrounded by questions relating to whether the event is notable enough for inclusion and whether enough time has passed for the event to be suitable for inclusion

in the encyclopaedia. There are still tensions between editors about when an event is too recent and whether it is notable enough for inclusion.

Secondly, the coverage of phenomena for which few (or no) published sources exist has elicited debate within the Wikipedia community in the past few years with a Wikimedia Foundation-funded project entitled the 'Oral Citations Project'. Started by a member of the Wikimedia Foundation Advisory Board, Achal Prabhala in 2011, the Oral Citations Project sought to employ Wikipedia editors in the capturing and publishing of sources of knowledge for which no published sources exist. Oral citations as conceptualised by the project leaders involved Wikipedia editors interviewing members of a community about a local phenomenon and then using the audio files from those interviews as citations in related articles. There has been heavy debate about oral citations because Wikipedia policy forbids what is called 'original research' defined as information for which 'no reliable, published sources exist' (Wikipedia:No original research) and a number of other subject areas will be implicated if oral citations are officially allowed. Despite this, proponents argue that Wikipedia must extend its idea of what can be included in the encyclopaedia if it is to truly become a home for 'the sum of all human knowledge' (Wikimedia Foundation, n.d.), as Wikipedia's motto indicates.

In Gieryn's terms, both breaking news events and oral citations represent boundary disputes in which groups of actors are struggling over the definitions of what Wikipedia is and what is should be. The outcomes of these disputes are to some extent determined by policies, but policies do not wholly explain why different outcomes occur and when policies are differently interpreted. Policies serve as tools for actors to use at the edges of the network, that is, at the article level, in determining how specific phenomena are represented. Despite the fact that policies aren't all-powerful, they are an important feature of boundary work because they articulate the hierarchy of values and principles

that are important to the community and determine the range of communicative actions and the vocabulary that can be used legitimately within Wikipedia.

4.5   Conclusion

The difference between what is possible and what an encyclopaedia should essentially cover catalyses numerous strains that find their way into debates at various levels of Wikipedia work as editors must decide what is essentially encyclopaedic and what can be included in the encyclopaedia without fundamentally changing what they believe Wikipedia *is*. Rather than debates about what Wikipedia should include or not include, these debates are also about how Wikipedia envisions itself as an encyclopaedia. As Wikipedians decide whether to employ old maps of encyclopaedic tradition or to break from tradition, community members are asking themselves what they can change and what they need to keep in order to retain their encyclopaedic status. Such decisions are not obvious nor are they inevitable; they represent the strains, interests and knowledge of those making decisions and the tools available with which actors can engage one another.

Verifiability is a fundamental principle of Wikipedia because it not only defines what can be included in Wikipedia, it also defines the roles and responsibilities of sources, editors and readers, as well as the relationship between them. The principle of verifiability signals a distinct departure from encyclopaedias of the past and has significant implications for what can be included in the encyclopaedia. With roots in the free and open source software movement as well as the scientific community, the verifiability principle defines the boundaries of communicative actions between readers and editors of Wikipedia. Wikipedia's particular instantiation of the logic of verifiability has

significant implications for what can be asked for and answered on Wikipedia and sets up a series of strains that must then be resolved at the article level.

In this chapter I have analysed the central philosophies and ideologies of Wikipedia that have been formulated against the backdrop of significant socio-technical change. Like changes in the ways that science was communicated in the past, we are seeing efforts to develop new centres of representative power by reaffirming the already legitimated power of academic professionals. Foundational to these trends is the verifiability ideology which plays a significant role in the way Wikipedia works and the types of knowledge that it represents. The verifiability policy determines what Wikipedia accepts and what it rejects, as well as how it defines who are the authors, subjects and experts. It sets up a particular vocabulary and series of communicative actions that can be taken within Wikipedia's socio-technical framework.

Wikipedia policies do not, however, ultimately define what specific claims will be accepted or rejected; they define a vocabulary and the strains that accompany that vocabulary. What is accepted or rejected on Wikipedia is dynamic and temporary, but every decision is significant. Decisions are dynamic and temporary because they are a result of the relative power of actors within the networks at the edges of Wikipedia's platform to have their definitions of Wikipedia's ideal experts, subjects and authors be accepted. Decisions are influential because every victory, albeit temporary, is used as evidence for further actions in boundary work later on.

In the next two chapters, we turn to two instances in which the principle of verifiability serves as the terrain in which facts travel through Wikipedia's socio-technical system.

Chapter 5: The slow progress of surr

> The "Encyclopaede" is like a Fordist factory of brutalisation and line work. And Wikipedia is like we've just taken over the factory for the workers… but the conveyer belts still discipline us. We accept the most heinous sources of knowledge on Wikipedia, as long as they're in a dead tree… or now represented at 1s and 0s. ([Fifelfoo] IRC interview, 9 February 2015)[20]

Surr is a sport played by people in the villages of northern India. The contested nature of the Wikipedia entry on surr can help us understand how facts travel through Wikipedia's infrastructure. Surr was one of the Wikipedia articles created for the 'Oral Citations Project'[21] conducted by Wikimedia Foundation Advisory Board member, Achal Prabhala, in 2011. The Oral Citations Project was initiated as a way of responding to the dearth of published material about topics of relevance to communities in the developing world.

In the opening to the project page about the project, Prabhala writes that the key problem that he is trying to address is the disparity between Wikipedia's global goals and the fact that printed and published knowledge is difficult to come by in many developing countries.[22]

> Imagine a world in which every single person on the planet is given free access to the sum of all human knowledge. To many within the Wikimedia movement, this idea is the guiding ambition that drives us. The problem with the sum of

---

[20] Written online sources have been reproduced verbatim throughout to reflect the original message including any errors, unless otherwise noted.

[21] Also known as the 'People are Knowledge' project.

[22] See https://meta.wikimedia.org/wiki/Research%3aOral_Citations.

human knowledge, however, is that it is far greater than the sum of printed knowledge. (Prabhala & Wikimedia authors, n.d.)

Prabhala writes that Wikipedia policies suggest that printed knowledge is privileged on Wikipedia but that books and printed material are a luxury of the 'rich economies (of) Europe, North America, and a small section of Asia.' He notes that there is very little scholarly publishing in languages other than English in India and that most South African languages other than English and Afrikaans have had a 'primarily oral existence'. Because of this disparity, the knowledge of some communities is privileged on Wikipedia, while others' knowledge remains unrepresented. The lack of written, printed and published material in countries and languages outside the Global North, writes Prabhala, is not only problematic for Wikipedias written in developing country languages; it is problematic for the rest of the world that this knowledge remains unrepresented.

As a result of this disparity, everyday, common knowledge - things that are known, observed and performed by millions of people - cannot enter Wikipedia as units of fact because they haven't been written down in a reliably published source. This means that not only do small-language Wikipedias in countries like India and South Africa lose out on opportunities for growth, so also does the Wikimedia movement as a whole lose out on the potential expansion of scope in every language. (Prabhala & Wikimedia authors, n.d.)

Prabhala suggested that a new type of source, an 'oral citation', be employed in cases where no published information about a topic exists. As part of the project, Prabhala visited communities where oral tradition has facilitated the transfer of knowledge about local customs and culture, interviewing residents about their activities as sources for articles on Wikipedia relating to those subjects. Members of the project conducted in-person interviews with community members in the rural village of Ga-Sebotlane in Limpopo province, South Africa and over the phone with interviewees from Kannur, a

city in North Kerala, India. Prabhala uploaded the audio files of the interviews onto Wikimedia Commons and asked Wikipedia editors to help write articles about their subject matter, using the oral citations as references (see table 5.1).

Table 5.1 Wikipedia articles created or added to as part of the Oral Citations Project

	Dappa Kali	Gilli-danda	Morula	Kgati	Neeliyar Bhagavathi	Surr	Tshere-Tshere
English WP	x	x			x	x	
Hindi WP		x				x	
Malayalam WP	x				x		
Sesotho WP			x	x			x

For the first five months after the articles were published, they remained unchallenged. In September 2011, however, all four English versions of the articles faced significant opposition as the implications of the project were surfaced on the Wikimedia mailing list and the details of the particular cases were debated on talk pages, noticeboards and other forums within Wikipedia.

In the next section, I provide a thick description of the surr article as it moved through Wikipedia's socio-technical structure, starting from its initial creation on Hindi Wikipedia in early 2011. For the proponents of the Oral Citations Project, the hope was that the fact would travel well within Wikipedia's infrastructure. Such travel could have resulted in the article being translated into other languages and attracting further information, photographs and perhaps even other sources, as well as being linked to new articles that might be created in the future about similar objects of indigenous knowledge inspired by the article.

At stake in the travel of the fact was not only how *far* along Wikipedia's network it would travel, but how the fact's 'subjects, experts and revolutionaries' (Jasanoff, 2004) would be defined by the actors who came into contact with it.[23] This chapter asks why surr did not travel far within Wikipedia's socio-technical system by applying Morgan's (2010b) framework regarding which factors are relevant to how far a fact will travel, and then asks how authority and expertise are being reconfigured as evidenced by the types of boundaries being set by actors in relation to a fact's traveling companions and Wikipedia's terrains and borders.

5.1 An account of the fact's travels

Wikipedia articles are constructed using numerous facts in a variety of forms. The primary fact is embedded in the article title that establishes the phenomenon (in this case, surr) as an 'autonomous, short, specific and reliable piece of knowledge' (Morgan, 2010, p. 8). Supporting facts include redirects to the article that establish facts about alternate names for the phenomenon, facts that make up the content of the article and facts that occur in categories and images, captions and infoboxes (as can be seen in figure 5.1)

The players of surr are the fact's subject, and a following of the fact involves following the actors, objects and the processes that become enrolled as the fact moved through Wikipedia's infrastructure. I followed surr from March 2011, when the article was first

---

[23] The distance that a fact travels in Wikipedia is defined by its translation into other languages on Wikipedia, being linked to other articles through interwiki links and links to other articles within the same category etc., links to other Wikimedia projects and links to sources outside of Wikipedia and Wikimedia (see Chapter 2.2).

created on Hindi Wikipedia, to August 2015, focusing particularly on the occasions when

surr and its traveling companions (notably its sources) were questioned and subjected

to rigorous debate.

The story of surr begins on Hindi Wikipedia in the country in which surr is played.

☐ Traveling companions
☐ Facts

Figure 5.1 The facts and traveling companions of surr

Source: https://en.wikipedia.org/w/index.php?title=Surr&oldid=674897959

5.1.1 Surr is first published on Hindi Wikipedia

On the 28th of March, 2011, an editor of Hindi Wikipedia, [SiddharthItripathi] published the first edition of the surr article (hi.wikipedia.org/surr). In the next hour, [SiddhartItripathi] added text to the article about surr's rules and a photograph of the field of play drawn onto a whiteboard, as well as a category titled 'Oral Citations Project'. In the next edit, [Mayur] added another category ('sports') using the HotCat tool.

HotCat is a MediaWiki gadget that enables registered users to easily remove, change and add categories to Wikipedia articles and MediaWiki is the free and open source wiki software on which Wikipedia is run. Gadgets are add-ons to the MediaWiki software that perform particular editing functions more efficiently. Such tools are regulated independently by those who design them. HotCat's information page on Wikipedia (Wikipedia:HotCat) indicates that it was initially developed for Wikimedia Commons by the user, [Magnus Manske] but was later adapted and improved for use on other Wikimedia projects by additional users (Wikipedia:HotCat). HotCat is only available to registered users and enables such users to easily remove, change and add categories to Wikipedia pages. Instead of a user seeing the categories as linked text, they become dynamic entities accompanied by functions that enable registered users to easily remove categories or propose new categories using an auto-completion feature that suggests categories to the user.

Hindi Wikipedia editors continued to make minor additions and changes to the surr article, adding references to the oral citations interviews hosted on Wikimedia Commons. The Wikimedia Commons pages for the citations contain audio files and

written transcripts created by [Aprabhala].[24] Each of the interviews has been conducted in Hindi, with Hindi transcripts provided for both interviews and an English translation available only for the first interview. Interviews were documented as taking place with members of the local community in the Rudrapur region of India and include an interview with a man who calls himself an 'Economic demonstrator for Uttarkhand Govt. in Rudrapur' according to the transcript (figure 5.2).

Figure 5.2: Interview with Mr Deepak Tripathi by [Siddharath]

Source: https://commons.wikimedia.org/wiki/File:PeopleAreKnowledge_Sur_Interview1.ogg (accessed 10 November 2014)

On the 14th of September, [Utcursch] moved the categories in the Hindi article to its standard position below the references and added a new image.

---

[24] I refer to Achal Prabhala as [Aprabhala] when referring to his editing activities and Prabhala when discussing his activities as a Wikimedia Foundation Advisory Board member.

5.1.2 Stabilizing surr on English Wikipedia

Edits 1-3:

[Utcursch] is an administrator of the English Wikipedia from India who was living in Canada at the time of the study. [Utcursch] made his first edit on Wikipedia in 2004 and became an administrator a year later. He has contributed to both Hindi and English Wikipedia, but prefers to edit in English because of the higher numbers of readers of the English Wikipedia. [Utcursch] also admits that he is 'put off' by the tendency of the Hindi Wikipedia community to use 'highly Sanskritized Hindi' ([Utcursch] interview, 24 November 2014). The official Sanskritized Hindi is seen by many in India as unintelligible to the general populace and used mostly in official communiqués and formal speeches (Kumar, 2013).

> My main motivation for contributing is the readership, not the zeal for language promotion and preservation. ([Utcursch] interview, 24 November 2014)

Like many administrators on English Wikipedia, [Utcursch] doesn't usually create new articles but spends the majority of his time rewriting articles and blocking vandals.

> Content creation is what I enjoy most, although I don't always get time for that. These days, I mostly work on cleaning up existing articles. Often, that involves a complete rewrite of the existing articles. My administrative work is mostly limited to blocking vandals, though I do end up performing administrative tasks related to the articles I'm working on (for example, deleting obsolete pages, merging article history etc.) ([Utcursch] interview, 24 November 2014)

In 2011, Prabhala reached out to a network of his contacts to find people interested in helping to write articles. He was aware that he did not have an extensive editing history under his [Aprabhala] account and that this may raise red flags when the articles were being assessed.

I had a personal, private editing account under which I notched up a fair number of edits btw 2004 and 2006, but I let it go soon after the advisory board of the wmf[25] was formed and I was asked to it – mainly because my work expanded and I no longer had the time to edit for fun… This meant that my editing history under the account linked to my name ([Aprabhala]) was not prolific, and I was nervous about creating articles from scratch – I thought it would attract undue attention, and I thought the article had a better chance of standing if stronger editors (not necessarily admins) did that instead. (Achal Prabhala interview, 5 March 2015)

[Utcursch] was one of the editors who had been contacted through Prabhala's network of volunteers and advisors to help with the article and it was [Utcursch] who translated the first version of the surr article. [Utcursch] translated the Hindi version of the article into English and added a category tag entitled 'traditional sports of India'. [Utcursch]'s use of categories linking this article to a series of other articles contained within 'traditional sports of India' enabled him to rhetorically link the claim to a series of other claims that had already been legitimated in other articles using the category.

[Utcursch] also replaced the photograph of the hand-drawn field of play in the Hindi version with a new image (see figure 5.3). In my interview with him, [Utcursch] noted that he had created the image using the free and open source software called Inkscape and saved it in the SVG (Scaleable Vector Graphics) format instead of JPEG (Joint Photographic Experts Group) produced by [Siddharth.tripathi] because the format was better suited to Wikipedia.

SVG images are better for many reasons, and are encouraged on Wikimedia Commons for diagrams, maps, charts etc. They are resolution-independent and they can be edited easily (e.g. if someone wants to replace the English text in image with German/Hindi text, this can be done in a normal text editor or an open-source program like Inkscape, without loss of quality). ([Utcursch] email interview, 6 March 2015)
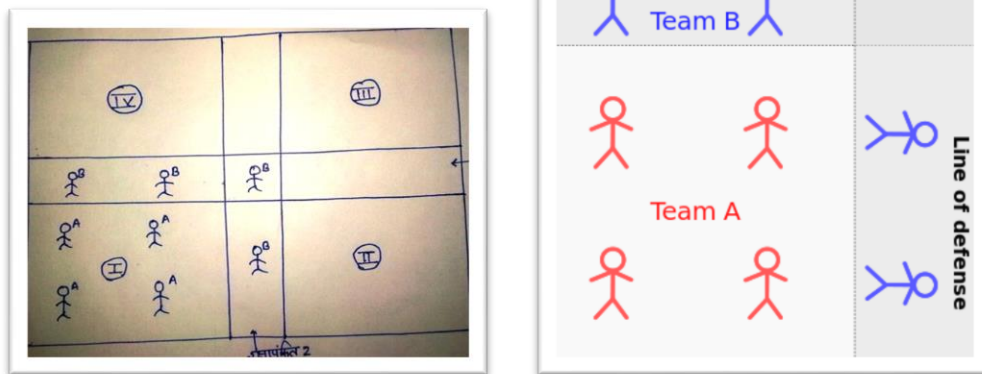
---

[25] The Wikimedia Foundation.

Figure 5.3 Surr images created by [Siddharth.tripathi] and [Utcursch] respectively

Source: https://commons.wikimedia.org/wiki/File:Surr.jpg and https://commons.wikimedia.org/wiki/File:Surr.svg

[Utcursch] also unconventionally added a hyperlink to the Oral Citations Project page that includes a description of the project's goals and methods. In Latour's (1987) terms, [Utcursch] was 'summoning allies' by adding the hyperlink to the oral citation project page, thus making explicit the connection of the articles to support from the Wikimedia Foundation. Then head of Global Development at the Wikimedia Foundation, Barry Newstead, as well as then head of fundraising, Zack Exley were listed as project contact and the Wikimedia Foundation contacts respectively (see figure 5.4). The summary also explicitly lists the project as hosted by the Wikimedia Foundation and includes 22 links to articles in newspapers and blogs mentioning the project (see figure 5.5).

With this single link to the project page for oral citations, [Utcursch] was engaging in a performative act that served to link the facts constructed by the project to high-ranking officials at the Wikimedia Foundation, the Wikimedia Foundation Advisory Board, professors and academics from three countries, a host of journalists and bloggers and

162

members of village communities in which the articles were based. This constituted a defensive move of shoring up support in case of attack.

Figure 5.4 Infobox for the Oral Citations Project hosted by the Wikimedia Foundation in 2011

Source: https://meta.wikimedia.org/wiki/Research:Oral_Citations (accessed 10 November 2014)

Figure 5.5 Screenshot from the Oral Citations ('People are Knowledge') Project page

Source: https://meta.wikimedia.org/wiki/Research:Oral_Citations (accessed 10 November 2014)

When [Utcursch] published the article at 7am UTC on the 14th of September 2011 and it was immediately available to read on the public Internet. [Utcursch] has autopatrol rights on Wikipedia which means that the article was not flagged in the 'New Pages Feed' where editors acting as page patrollers evaluate new articles. As an autopatrolled user, [Utcursch]'s contributions are automatically marked as patrolled by the MediaWiki software. According to the Wikipedia page relating to the autopatrol process,

> The autopatrolled… user right is intended to reduce the workload of new page patrollers and… it means that the user can be trusted not to submit inappropriate material, deliberately or otherwise. (Wikipedia:Autopatrolled)

Users must apply for autopatrol rights but administrators are automatically granted autopatrol permissions. The allocation of autopatrol rights effectively means that new articles created by users in this category are not subject to the same level of scrutiny in the new pages feed (although they may still come under scrutiny through other processes as happened later in this case).

Edit 4-6:

Ten days after the article was published, [Aprabhala] improved the grammar in the article by replacing a few words in the main text and restructuring sentences. In early February the next year (2012), an anonymous editor with an IP address that geolocates to South Korea added an interwiki link to surr in Korean. Interwiki links are links on the left-hand side of a Wikipedia page that provide links to other language versions representing the same fact or phenomenon. Figure 5.6 provides an illustration of the first six edits of the surr article.
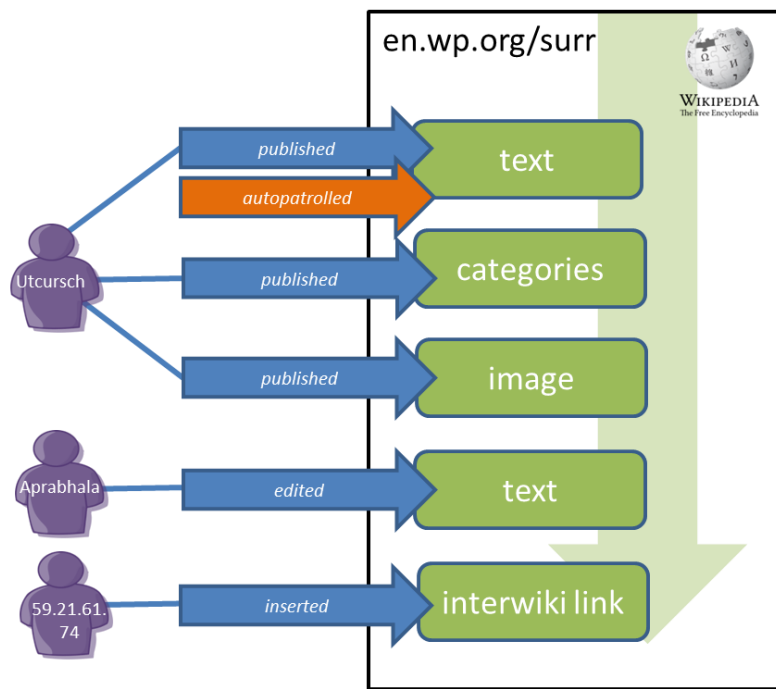
Figure 5.6 The consolidation of surr on English Wikipedia

Source: author

5.1.3 The article (and those associated to it) are challenged

A few days later, [Around The Globe] posted a question to the Reliable Sources Noticeboard (RS/N) on English Wikipedia (Wikipedia:Reliable Sources/Noticeboard) about the citations used in the surr article. According to [Fifelfoo] who is a frequent contributor, RS/N is where editors come to obtain advice about whether a source is reliable or not according to the interpretation of Wikipedia's reliable sources and verifiability policies discussed in the previous chapter. [Fifelfoo] highlights the wide interpretability of Wikipedia's foundational policies regarding sources and the need for more contextual information about particular decisions.

> The RS/N guideline is *Identifying reliable sources*, the policy is *Verifiability*. But you can drive a truck through those. So RS/N is kind of a vibe of "common law" findings. ([Fifelfoo] IRC interview, 9 February 2015)

[Fifelfoo] described himself in my interview with him as 'a 36 year-old male, born, raised and living in Australia' who has been editing Wikipedia since 2003. [Fifelfoo] works for an Australian university in administration and labour history research and is finishing his PhD in another university. [Fifelfoo]'s user page contains numerous requests from other editors for his advice on questions about reliable sources indicating that he is viewed as an expert on Wikipedia sources by many other editors. [Fifelfoo] believes that he gained this reputation through his contributions to an essay offering advice on examples of reliable sources for history articles and by being vocal about the need for 'well researched and consistent citations' (IRC interview, 9 February 2015). Through this, [Fifelfoo] believes that he became 'known for knowing... how to read knowledge from texts, and (how to) cite knowledge.' (IRC interview, 9 February 2015)

In this case, twelve editors (including [Fifelfoo]) debated whether oral citations constituted reliable or unreliable sources on English Wikipedia. This debate constituted an exercise in the classification of the novel oral citations sources as either unreliable or reliable according to the ideological strains and interests of the editors involved. Like Gieryn's scientists excluding rivals from within by defining them as outsiders, the majority of editors in this discussion used disparaging adjectives to label the authors of the oral citations as 'unreliable' and 'amateur' as opposed to 'professional' or 'academic', and the citations themselves representing 'opinion' rather than 'fact'.

The excerpt below demonstrates how the interviewee of one of the oral citations was labeled as a 'layperson' (as opposed to an expert) and the fact's notability as 'unknown', thus casting doubt on its suitability for Wikipedia.

(T)he person interviewed has no academic authority in the field. He is simply a layperson, who has played a game of unknown notability. ([NativeForeigner], RS/N discussion, 13 February 2012)[26]

In another statement, the publisher of the source ([Aprabhala]) is labeled as a 'commons content creator' as opposed to a 'professional or academic ethnographer'.

I'm sorry, but commons user Aprabhala is not a professional or academic ethnographer; they're a commons content creator. ([Fifelfoo], RS/N discussion, 13 February 2012)

In the statement below, [Luke Warmwater101] labels oral citations as 'primary, or even self-published sources' (which are cautioned against in Wikipedia policy), and as 'opinion' rather than 'fact'.

In general, oral citations cannot be checked for accuracy. They are therefore akin to primary, or even self-published sources. Because of that, at best they might be used as statements of opinion, but not as statements of fact. ([Luke Warmwater101], RS/N discussion, 13 February 2012)

Similarly, [NativeForeigner] suggests the same definition of oral citations as 'good enough' for the article but as not sufficient for providing notability.

My inclination is that they aren't good enough for notability, but are good enough to be used as sources. Even in academic research audio sources are acceptable: notability though is an entirely separate matter, and must be decided separately. ([NativeForeigner], RS/N discussion, 13 February 2012)

None of the editors participating in the discussion on the RS/N articulated alternative definitions of the oral citations sources, focusing instead on re-defining the problem of the lacuna in published sources from many parts of the world that catalyzed the project in the first place and suggesting alternative solutions. In one exchange, [Smallbones] asked what is the solution to representing local knowledge for those who do not have a

---

[26] The RS/N discussion about oral citations is archived at https://en.wikipedia.org/wiki/Wikipedia:Reliable_sources/Noticeboard/Archive_115#Oral_Citations.

rich publishing tradition. S/he writes that the solution to translate articles from those

with extensive published sources would constitute 'cultural imperialism'.

> I do suggest that everybody actually read the research. I don't think it applies especially well to the English language Wikipedia. But I do think we should address the question posed by that paper. There are 200+ language versions of Wikipedia. Perhaps 50 of them have extensive book, newspaper, and magazine publishing - things we'd accept as reliable sources. So what are the 150+ left out languages supposed to do? Just translate from the 50 "published languages" - that would indeed be cultural imperialism. Say that they don't have, in their own terms, reliable sources - ditto. So do you have an answer for them? Don't think it matters? - ditto. ([Smallbones], RS/N, 14 February 2012)

[Fifelfoo] responds that smaller language Wikipedias should develop standards and a systematic method for ethically collecting oral records.

> If they want to use a shitty epistemology/methodology like the Dutch encyclopaedia, then I will frown upon them and others may follow. If they want to use a less stringent epistemology/methodology like the French or Germans, then I will be less pleased with their work and others may follow. If they wish to meet the epistemic and methodological standards of the en. culture, while making use of repositories of oral knowledge, then they will need to develop a systematic method to ethically collect and copyright release oral records, and a method to utter them in the sense of publication in a mode that demonstrates standards equivalent to the demands of reliable sourcing... ([Fifelfoo], RS/N, 14 February 2012)

Editors continued to discuss alternative solutions, among them the suggestion that an

additional institutional structure be created, separate from Wikipedia, where experts

with 'certified professional capacities' ([Fifelfoo]) or 'widely acknowledged authorities

in the field in question' ([Mangoe]) publish knowledge relevant to local contexts. Editors

differed as to how such independence might be achieved. For [Mangoe], independence

would be signaled by the existence of 'widely acknowledged authorities in the field in

question' – that is, expertise that is recognised outside of Wikipedia. For [Fifelfoo], on

the other hand, independence could be achieved with a 'closed wiki' where certified

professionals could authorize publication of works as reliable.

Edit 7

Catalysed by the RS/N discussion, [Fifelfoo] removed the hyperlinks to the oral citations in the reference list of the surr article on English Wikipedia so that they were no longer linked to the interviews. He also changed the wording of the citations noting that the interview was 'in Hindi' and 'possibly with Siddharth Tripathy'. [Fifeloo] then added a warning tag to the interview on Wikimedia Commons indicating that it if evidence of copyright permission was not received within seven days, the file would be deleted.

[Fifelfoo] had stated on the RS/N discussion that there was no indication in the upload record of the oral citations that the respondents had released their copyright and that this made the interview an 'obvious' violation of the respondent's copyright ([Fifelfoo], RS/N, 13 February 2012). When editors upload media to Wikimedia Commons, they must typically provide a copyright waiver form in which it is stated that the author/s of the work dedicate their content to the public domain or license it under a Creative Commons Attribution Share-Alike license.

In this case the interviewer asked for permission from the interviewee for their conversation to 'be kept in Wikipedia Commons' to which Deepak replied, 'I have no objection with that' (Deepak Tripathi, Wikimedia Commons, 27 June 2011). [Fifelfoo] complained that the interviewee's copyright had been violated because there was no copyright release form accompanying the interview. [Nearly Headless Nick] (also known as [Sir Nicholas de Mimsy-Porpington]), a native Hindi speaker and administrator of English Wikipedia, responded that the interviewee had said in Hindi 'that they do not have any issues with their work being published in Wikipedia' but noted that this statement was 'neither specific enough nor in writing' ([Nearly Headless Nick], RS/N, 13 February 2012).

As a corollary to obtaining access to allies, cutting off allies is a strategy for destabilizing the article. In this example, [Fifelfoo] removed the hyperlink to the source in both citations, marking his edit with a comment indicating that the content had been removed ('snipped') because of a copyright violation. By removing the link to both the source and the project page, [Fifelfoo] was effectively cutting the author and his claims off from his allies using the technical vocabulary of Wikipedia's copyright management. By removing the citations from the article and tagging the interviews as missing copyright permissions, [Fifelfoo] was effectively defining the fact as unreliable and unstable.

When I asked [Fifelfoo] in my interview with him whether his main concern with the project was copyright, he responded:

> No, copyright was a stick to hit them with because they didn't have appropriate permissions statements and had a naïve attitude towards knowledge in communities and power. ([Fifelfoo], IRC interview, 9 February 2015)

[Fifelfoo] believed that the participants of the Oral Citations Project were abusing their power in obtaining audio interviews for an essentially Western project from the hapless individuals.

> Now I don't have direct experience of non-literate communities. But I've heard enough from people representing communities with oral traditions about the importance of other people not misusing the knowledge in third spaces; for example, by publishing travelogues, or unethical academic papers.

> And one experience which has stuck in my life was the poor quality of papers written about online communities in the 1990s by authors who didn't even bother with the MLA's standards for research in complex texts, let alone getting human ethics clearance.

> So I normally view one motivation to privilege oral accounts as a kind of "Orientalism" combined with a ram-raiding research vibe of smash down the cultural barriers, grab all the nice ideas, and fence them for kudos. ([Fifelfoo] interview, 9 February 2015)

[Fifelfoo] was effectively employing the copyright tool to oppose what he believed was an instance of 'ramrodding' indigenous communities (interview, 9 February 2015).

Edit 8:

Without links to its allies (or any traveling companions in the form of citations), the fact became vulnerable to further attack which occurred when, in the next edit, [Sir Nicholas de Mimsy-Porpington] added two warning tags to the head of the page. The tags indicated that the article needed citations for verification and that it relied on primary sources, requesting editors to add references to secondary or tertiary sources. This action served, once again, to label the facts as unreliable and unstable by defining sources as primary and therefore not as acceptable according to Wikipedia's reliable source policies.

A week later, the Oral Citations Project was discussed on the Wikimedia-l mailing list. Wikimedia-l, is one of about 400 mailing lists administered by the Wikimedia Foundation but it occupies a special position in the Wikimedia network because of its role in providing a forum for issues relating to all Wikimedia Foundation projects. The Wikimedia-l mailing list is used to discuss issues relating to new projects, new chapters, polling and fundraising and is where the majority of high-level, strategic and long-time discussions about movement-wide issues take place between the Wikimedia Foundation and non-Foundation members of the Wikimedia community.

The discussion about oral citations began when Prabhala sought advice from the Wikimedia community about the Oral Citations Project.[27] During the discussion,

---

[27] Archived discussion available at https://lists.wikimedia.org/pipermail/wikimedia-l/2012-February/thread.html.

Prabhala noted that it was problematic that Wikipedia policies left out knowledge held by so many people. He pointed to the URL of the Oral Citations Project[28] that was being hosted on meta-wiki, a site for coordination, planning and documentation of Wikimedia projects.

In response, fourteen editors weighed in on the question. There were three key disagreements among those responding to the post on the mailing list. They included disagreements about Wikipedia's encyclopaedic identity, about the role of Wikipedians in terms of the knowledge that they represent, and about the subjects of knowledge and the role of those subjects in the representation process.

Firstly, editors disagreed as to whether Wikipedia's identity as an encyclopaedia compelled them to either adhere to encyclopaedic traditions by relying solely on academic experts to record knowledge or whether, as a digital encyclopaedia written by amateurs, they could depart from classical encyclopaedias by producing knowledge themselves. Former Wikimedia Foundation board chairperson, Ting Chen, argued that the Oral Citations Project signalled an opportunity for Wikipedia to do things differently from traditional encyclopaedias. Chen told the story of the first editor of Chinese Wikipedia, [Mountain], who wanted to record the traditional practices of local fishermen but was prevented from doing so because there were no academic sources available on the subject.

> He came from the coast of Shandong, and his father told him that earlier there was a local tradition where people went early morning to the coast to catch crabs or mollusks (one of them). They used to use a special technique to catch the animals. But meanwhile no one is using this technique anymore, not only because there are now plenty of crabs or mollusks on the market from the hydroculture, but also because the coast which was wild earlier are now all

---

[28] See https://meta.wikimedia.org/wiki/Research:Oral_Citations.

urbanized, with oil terminals and harbors and those. When Mountain told me that story he felt he would like to write down those stories because in maybe 10 or 20 years, latest in 50 years, no one would ever know that there was such a thing on the world. And that tradition would be lost [forever]. But he also felt he could not write them on Wikipedia because he had no resources, because until now no of the (ethnologists) ever had interested on such traditions and no academic resources ever mentioned it. With the Oral Citations Sourcing it would be possible to interview the old people or even let them show how the techniques worked. (Ting Chen, Wikimedia-l mailing list, 25 February 2012)

Another editor, Ziko van Dijk responded to Chen by arguing that only academic experts could legitimately record knowledge held by 'individuals' or 'witnesses':

Yes Ting, and for these cases there is the method of [[oral history]].

This is a means to create what the Anglosaxons call "primary sources". It is recorded and can later be used by a scholar (historian, ethnologist etc.) for his research, for his "secondary sources". These, with their scholar reflections, can be used by an encyclopedia.

There are good reasons for this way. One is, that it is not very practical to cite from audiotapes/audiofiles. Another, that what this individual is describing may be true for his personal environment but cannot be generalized to others. For that, one needs the scholar. Remember: witnesses are the most unreliable source ever. People tell you plain nonsense - not because they want to (lie) or are stupid but because the human brain is simply not created to be a historian. It has the greatest difficulties to store information truthfully. So you need to record, and compare the different assertions from different people.

It is a possibility to record oral and visual expressions from illiterates, and only later to do something with it scholarly. But all this has nothing to do with Wikipedia. (Ziko van Dijk, Wikimedia-l mailing list, 25 February 2012)

Chen responded that van Dijk's was the traditional view of how encyclopaedia should be produced, but that Wikipedia could produce the encyclopaedia differently; Wikipedia, *was* in fact, already doing things differently:

Yes, it is the way how classic encyclopedia worked. But Wikipedia is not a classic encyclopedia, and I don't see the sense to bound ourselves possibilities just to please some old traditional rules.

Classic encyclopedias were written by scholars, Wikipedia is not. Wikipedia say, everyone can work on an encyclopedia, and because of this, the content included in Wikipedia is far more richer and broader than in a classic encyclopedia.

Scholars have limited capacities. A lot of things scholars cannot pay attention to everything. In give everyone the possibility to pay attention to what they think is

interesting and important in their life, we can free a lot of potentials that the scholars cannot.

Scholars have their own point of view. Over eons scholars thought only what kings and knights did are important for the history. We know today this is wrong, and the scholars of today are happy if they can find a garbage dump from a village of thousands of years ago, so that they can catch a [glimpse] about how people lived then. Suppose the people in those villages could have recorded their lives with the oral citations.

In the history of science it is proven many times that also amateurs can provide breakthroughs and insights that scholars neglected. Done in an orderly way oral citations on Wikimedia projects can provide a lot of sources and informations that the scholarship are not able to because of the lack of capacity, interest, ideology. (Ting Chen, Wikimedia-l mailing list, 25 February 2012)

This engagement demonstrates the epistemological questions being asked as a result of the Oral Citations Project. Van Dijk argued that the knowledge that Wikipedia represents should be recorded by experts. He defined experts as 'historians' and 'ethnologists' (that is, professionals from the academic field) and distinguished them from 'individuals' and 'witnesses' who van Dijk argued were 'unreliable'. Chen, on the other hand, denigrated academics as ideal recorders of the knowledge held by ordinary people by arguing that, in the past, academics have either ignored certain subjects because they can't cover 'everything' or have covered subjects in a biased manner by focusing their recording on the stories of the powerful ('kings and knights') rather than ordinary people. Whereas van Dijk argued that 'individuals' are 'unreliable', Chen proposed that Wikipedia has demonstrated that 'amateurs' can do what academics cannot or will not, and that this represents an opportunity to record knowledge that has until this time been unrepresented.

Van Dijk argued that Wikipedia should not be in the business of recording knowledge in published form, and that 'this has nothing to do with Wikipedia'. If subjects of knowledge want to have their knowledge on Wikipedia represented, then they need to

create sources so that Wikipedians can cite those sources in encyclopaedic representations of them.

> Those people who would like to write on Wikipedia about any subject can write a book or pdf about it. It does not have to be a scholarly work in every aspect. And then, the Wikipedia in language X can decide that it accepts this kind of literature as reliable. (Those various standards are not uncommon in the different Wikipedias.)
>
> Not everything has to happen *in* Wikipedia. (Ziko van Dijk, Wikimedia-l mailing list, 25 February 2012)

Castelo responded that those with knowledge subjects, those with the expertise and knowledge relevant to Wikipedia, are sometimes not able or willing to produce academic texts about their knowledge. Speaking about his experience with indigenous knowledge in Brazil, Castelo wrote:

> … the people who tells the facts are not the same people who want to write on Wikipedia, and definitively, not people willing to write a book or pdf… Some of the indians cannot write a book, a pdf or a Wikipedia article and those are exactly who have more expertise on their traditions. This can give them authority when describing their rituals, clothings, artefacts, fights, cuisine, etc., much more than a wikipedian can. And we still have a huge lack on articles about them, because for certains indigenous nations, there are almost no published material (some have no written material at all, as far as i know). I live in the capital city, where some of them usually come for present their culture in a national museum, and go back to their territories. In moments like these, we wikimedians can go there, take photos and record an interview (most speak a bit Portuguese, as well as their own languages), for publishing in Commons and Commons/Wikinews, respectively, for using in Wikipedia articles. (Castelo, Wikimedia-l mailing list, 25 February 2012)

Castelo argued that the knowledge subjects (in this case, indigenous people of Brazil who have knowledge about local culture) are the authorities rather than Wikipedians themselves. By designating the subjects of knowledge as the authorities rather than the recorders of knowledge, Castelo is making an argument that the recording of such authority is less important than the fact that authoritative knowledge in a variety of forms is represented on Wikipedia. The argument, in other words, is an argument about which identities are considered authoritative to different groups and what the basis of that authority is, whether it is the experience of a phenomenon or the accurate

recording of that phenomenon. Castelo, in this case, believed that authority should be extended to those with experiential knowledge of a phenomenon, rather than only the recorders of this knowledge.

As Ray Saintonge proposed in his reply below, the question of authority is not one that can be answered by Wikipedians themselves. Authority is decided by readers as well as editors. The question about whether Wikipedians should create primary sources is about whether Wikipedians should interpret those sources for their readers and this cannot simply be answered by Wikipedians alone.

> Waiting for historians or other "experts" to provide reliable secondary sources simply won't work. What these experts choose to write about is just as much a matter of whim as a volunteer's choice of what to add to Wikipedia.
>
> Oral citations are only one facet of the challenge. Wikipedia is only one facet of the solution. Restructuring the epistemology requires a larger vision than the one that gave us Wikipedia. (Ray Saintonge, Wikimedia-l mailing list, 27 February 2012)

The discussion about oral citations ended with another comment from Saintonge after less than a week of replies. When the discussion ended there was no clear consensus on whether allowing oral citations as defined by Prabhala and his colleagues was conducive to Wikipedia's goals or identity. None of the editors participating in the thread on Wikimedia-l edited the article during or after this discussion, but work continued on the article by a variety of actors and actants engaged in the daily practice of article maintenance and construction.

5.1.4   Bots enter the scene

Edit 9:

On the 24th of February, while the discussion on Wikimedia-l was still taking place, the two warning tags on the article were replaced by one consolidated tag by the bot, [[Battybot]].29 The book icons were replaced by the exclamation mark '!' alert with the line: 'This article has multiple issues. Please help improve it or discuss these issues on the talk page', followed by the alerts to issues regarding citations and primary sources. The difference between the two edits can be seen in figures 5.9 and 5.10 below.

Bots such as [[BattyBot]] are automated agents that are defined by Wikipedia as 'programs or scripts that make automated edits without the necessity of human decision-making' (meta/Bot policy). Bots are distinguished from 'assisted editing' tools that leave ultimate decision-making up to the user. Whereas assisted editing tools are characterised as 'low-speed tools and scripts' that assist users to make decisions, bots make editing decisions according to a set of rules (that is, according to their algorithm).

A group of Wikipedia volunteers called the 'Bot Approval Group' (BAG) must approve each bot task, typically after a short trial during which time the bot is monitored and community feedback responded to. In this case, [[BattyBot]] is operated by [Goingbatty], who is listed on her/his user page as one of the most 'active Wikipedians of all time' with over 10,000 edits (en.wikipedia.org/User:GoingBatty). The task that [[BattyBot]] performed in the surr article was to merge a list of warning tags into a single {{multiple}} issues' tag. This task is listed on [[BattyBot]]'s user page as 'approved' but 'inactive'. The reason provided is that the {{multiple issues}} template was changed and that it has become more challenging to create the list within the template.

---

29 Bot names are indicated using [[]]s.

[[BattyBot]]'s edits (consolidating the two tags published by [Sir Nicholas de Mimsy-Porpington]) are understood by the Bot Approval Group as non-controversial and thus in keeping with bot policy, but the application of a new tag presents new information that could influence both the perception of readers and the management by editors in relation to the content that it frames. Instead of criticisms relating to the need for 'additional citations' and the article 'rely(ing) too much on… primary sources', the bot automatically consolidated these into a new statement about the content below it, that is, that the article had '*multiple* issues' (my emphasis) (see figure 5.7 and 5.8 below).
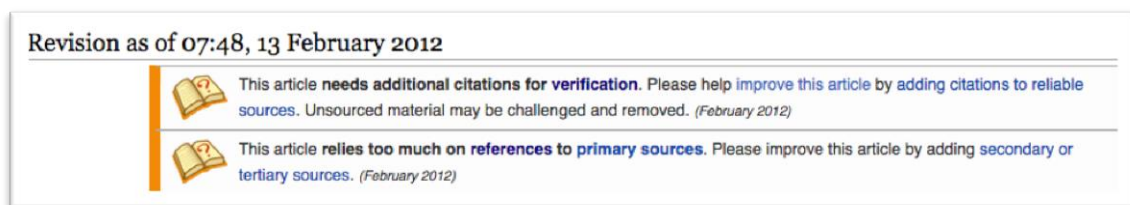
Figure 5.7 Revision as of 13 February 2012

Source: https://en.wikipedia.org/w/index.php?title=Surr&diff=prev&oldid=476608188
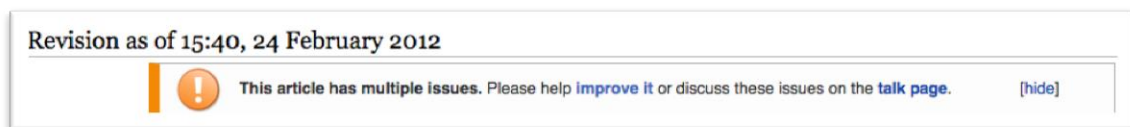
Figure 5.8 Revision as of 24 February 2012

Source: https://en.wikipedia.org/w/index.php?title=Surr&diff=next&oldid=476608188

5.1.5   A period of instability

Edit 10:

On the 7th of March, the entire text of the article except for the first line and the illustration with its caption was removed by [MER-C], an administrator on the English Wikipedia and a member of Wikiproject Spam and Wikiproject Copyright Cleanup groups who has been working on Wikipedia since 2006 (see figure 5.11 below). [MER-C]

178

summarised his/her edits with the explanation that the article was unsourced and that it failed Wikipedia's verifiability and reliable sources requirements.

Figure 5.9 Revision as of 7 March 2012 by [MER-C]

Source: https://en.wikipedia.org/w/index.php?title=Surr&diff=next&oldid=476608188

Edit 11:

Seven months later, on the 23rd of October 2012, [Utcursch] removed the warning tag and restored the text with the edit summary 'see meta:Research:Oral Citations' that was linked to the project page for oral citations. In restoring the text and referring, once again, to the project page for oral citations on meta, [Utcursch] was responding to [MER-C] by reminding her/him of surr's allies mentioned earlier. In my interview with him, [Utcursch] stated that he

> restored the content on the grounds that the oral citations were the result of a Wikimedia project, and should be considered reliable. ([Utcursch] interview, 24 November 2014)

In the next edit, however, [Utcursch] added a citation to an online news story in Hindi[30] two minutes later. The Jagran.com article is in Hindi which is not ideal according to policy, but it is at least published independently from Wikimedia in a newspaper that is widely read in India. According to [Utcursch],

> When other editors tagged the article as poorly sourced, I started searching for other sources that would corroborate the existence of this sport. I couldn't find any English language sources, so I added the only article I could - a Hindi newspaper article. It was just a way of saying "*This isn't completely made up! Here's a reference that can be considered as reliable by traditional Wikipedia standards.*" ([Utcursch] interview, 24 November 2014)

By adding the source, [Utcursch] was essentially complying with the definitions set by previous editors of what constituted a reliable source on English Wikipedia.

Edit 14 & 15:

On the 22nd of March the next year (2013), the two interwiki links were automatically migrated to Wikidata by [[Addbot]] so that they were no longer editable using the article editing interface but transcluded and editable at Wikidata.org. Transclusion is a process by which the content of one document is included in another document by reference. Wikipedia's MediaWiki software enables the content of a template to be included in a target page so that changes to the template are then automatically reflected on all pages into which that page has been transcluded.

On the 8th of May, the article was vandalised by an anonymous editor, [64.222.71.242][31] (IP address resolves to Indiana, USA), who deleted all the content and replaced it with

---

[30] The link was dead and the website had been taken offline at the time of this study.

[31] Anonymous editors on Wikipedia are those who do not log in to the site when they edit an article. Their editors are indicated by the IP address that they logged in using and the IP address can be geolocated in order to see which country and region the edit came from.

what appears to be humorous text as seen in figure 5.12 below. After [[Cluebot NG]] automatically reverted the article back to the version before the anonymous user edited it, it was then vandalised again five minutes later by the same anonymous editor.

> **Revision as of 18:27, 8 May 2013**
> He is a surr-dom and has a behhen attitude. he just surrs it all the time and i like to see him all behhed up and surred. HE is surred of life. ha.hah.ha. quite surrin me.....!
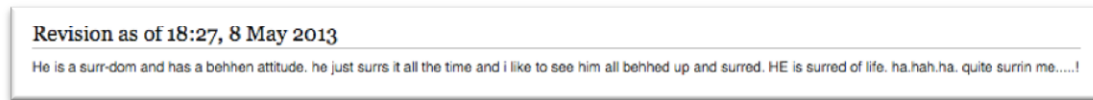
Figure 5.10 Revision as of 18:27, 8 May 2013

Source: https://en.wikipedia.org/w/index.php?title=Surr&oldid=554166687

Edit 16:

In both cases the vandalism was removed within seconds of being created – the first by the anti-vandalism bot, [[Cluebot NG]], the second by a member of Wikipedia's counter-vandalism unit, [Josve05a]. Whereas [[Cluebot NG]] automatically evaluated the edit and categorised it as vandalism before deleting it, [Josve05a] used the anti-vandalism software, Huggle, in order to detect and remove the content.

Huggle is an external browser application for dealing with Wikipedia vandalism that transforms, contextualises and abstracts a feed of recent changes to Wikipedia by applying a series of heuristics (see figure 5.11). Huggle selects which edits to show to vandal fighters according to variables that indicate whether edits are by new users, for example. The tool enables users to edit articles without a deep understanding of particular subject areas because they only have to answer the question: 'Is this edit a constructive or destructive edit?'
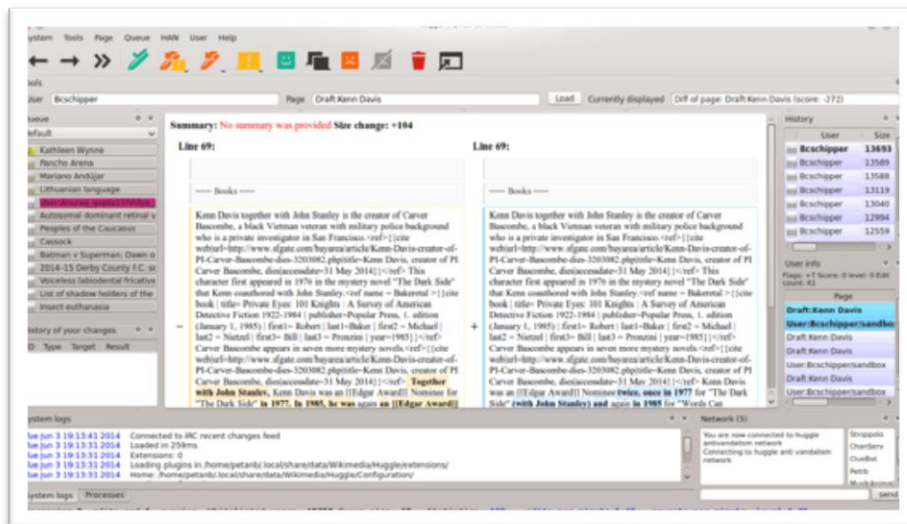
Figure 5.11 Screenshot of the Huggle interface

Source: https://en.wikipedia.org/wiki/Wikipedia:Huggle

Huggle's interface is designed to aid rapid decision-making for vandalism control by highlighting only the relevant text that has been added and by comparing that text to the text of the previous edit in the view beside it. Keyboard shortcuts enable the space bar, for example, to be used to progress to the next edit. On the top right of the interface is a score provided for the likelihood that the edit is vandalism and a series of actions that the user can take in order to warn the user and revert their edit, or to report the editor. The warning will automatically result in a notice being added to the offending editor's user page remarking that they have performed a destructive edit.

Huggle is an almost-daily site of practice for [Josve05a], a 16-year old Wikipedia editor from Sweden who has been editing Wikipedia for about four years. Asked what a day in the life of a vandal fighter was like in an interview, [Josve05]a responded, '(S)ome hate messages, spam, lot of laughter and a few hurt feelings' (IRC interview, 24 November 2014). Apart from anti-vandalism on Huggle, [Josve05a] also reviews files on Wikimedia Commons that require license review as well as reviewing new articles from anonymous

users. S/he also frequently uses the IRC channel, #wikipedia-en-help to help users, especially people representing companies who want their own page but who have had their articles removed. The IRC channel enables those with questions to be able to chat with a volunteer editor synchronously, that is, in real time. [Josve05a] remarked that [[Cluebot NG]] was the 'Best thing for anti-vandals, since... ice cream' (IRC interview, 24 November 2014).

When I asked her/him whether s/he still had a job removing vandalism when [[Cluebot NG]] was so effective at removing vandalism automatically, s/he wrote that the bot was often unavailable and that it reverts the same user on the same page only once every 24 hours and only if it is above 85% certain that it is vandalism. This explains why [[Cluebot NG]] caught the first instance of vandalism but [Josve05a] was able to catch the second.

Edit 19:

In February 2015, [Fifelfoo] removed the links to the oral citations, with a reference to the discussion on the RS/N in which the oral citations were discussed.

> Per: RSN https://en.wikipedia.org/wiki/Wikipedia:Reliable_sources /Noticeboard/Archive_115#Oral_Citations meta is not en ([Fifelfoo], edit summary, 3 February 2015)

[Fifelfoo]'s edit summary indicated that 'meta is not en' (meta is not English Wikipedia), thus denying the legitimacy of the project page for oral citations on meta as relevant to English Wikipedia and instead pointing to the archive of the RS/N as the authoritative venue for discussions relating to the article.

Six weeks later, [Matthewvetter] reverted [Fifelfoo]'s edit with his own edit summary that stated 'See Talk page for rationale for inclusion of primary/oral research for this

article' ([Matthewvetter], 17 March 2015). [Matthewvetter] indicated why he had reverted [Fifelfoo]'s edit on the talk page of the surr article.

> I reverted the most recent censure of oral citations made by Fifelfoo. I understand the issue of primary research, but in this specific case, because the subject cannot be represented well with secondary research due to the orality of the knowledge culture, primary research is appropriate. Per guidelines on "No Original Research," "A primary source may only be used on Wikipedia to make straightforward, descriptive statements of facts that can be verified by any educated person with access to the primary source but without further, specialized knowledge." ([Matthewvetter], en.wikipedia.org/surr talk page, 17 March 2015)

[Matthewvetter]'s definition of the relevant authority and authoritative space for conversations have remained unchallenged at the time of this study (1 August 2015). [Fifelfoo] is no longer active on Wikipedia (his last edit was the edit to the surr article on the 5th of February 2015), [Sir Nicholas] has only performed about 20 edits in since January 2015 and [Aprabhala] has made no article edits since 2012.

5.2   How well did the fact travel?

The surr fact did not travel fruitfully or with great integrity within Wikipedia's socio-technical system in the past four years. In terms of its integrity, there are conflicting statements about who the players of surr are so that the fact is accompanied by traveling companions that do not support its representation in the article. Although the text of the article indicates that surr is 'usually played by men', for example, the interview with Deepak Tripathi[32] indicates that the game is played 'by children' and the second interview with Siddharth Trivedi makes no reference to men in the interview transcript.

---

[32] See https://commons.wikimedia.org/wiki/File:PeopleAreKnowledge_Sur_Interview1.ogg.

In terms of fruitful travel, surr has only been translated into three languages and it contains few links to and from it. It is common to find articles in only one or a few languages across the different language Wikipedia. Hecht and Gergle (2010), for example, studied 25 different Wikipedia versions and found that 74% of articles were described in only one language, and 95% appeared in fewer than seven languages. In terms of in and out links, only one other English Wikipedia page links to surr (a list of articles with 'sur' in the title)[33] and it contains only one link to another Wikipedia page in the category of articles called 'Traditional sports of India'[34]. The lack of links to and from Wikipedia articles is discouraged in terms of policy, with the Wikipedia 'Manual of Style' noting the importance of linking articles to one another in an effort to aid the reader in finding relevant articles

> Linking through hyperlinks is an important feature of Wikipedia. Internal links are used to bind the project together into an interconnected whole. Interwiki links bind the project to sister projects such as Wikisource, Wiktionary, and Wikipedia in other languages, and external links bind Wikipedia to the World Wide Web. (Wikipedia:Wikipedia Manual of Style)

Orphaned (also known as 'lonely') pages are pages where there are no incoming links from other Wikipedia pages, that is 'pages with no parent' and there is a wikiproject called 'Project Orphanage' to clear up the backlog of orphaned articles. According to the talk page for the list of 'orphaned pages', the lack of incoming links could indicate 'that the article was created in error, that the articles which should link to the orphaned article need expansion or that the article is not noteworthy'. Without being referred to by other pages, an article tagged as an orphan could be seen as not noteworthy and deleted.

---

[33] See https://en.wikipedia.org/wiki/SUR.

[34] See https://en.wikipedia.org/wiki/Category:Traditional_sports_of_India.

Surr was accompanied by actions taken by those who were attempting to consolidate the fact and catalyse its travel within Wikipedia vs. those who were attempting to discredit and cast doubt on it. Editors employed both descriptive and performative tactics in order to either shore up support for the fact or to discredit it. Tactics were performed according to the norms and affordances of the spaces in which discussions were taking place (see table 5.2).

Discussion spaces such as Wikimedia-l, meta and the RS/N enabled editors to use rhetoric in order to discredit or add support to the fact. Remote editing through Wikimedia and Huggle took place using the tools and heuristics that govern those spaces. Automated editing by bots was initiated according to the rules designated by the bot operators and governed by the Bot Approval Group, but at a higher scale than human editors could accomplish and well removed from the context of surr since rules for bots are designed in advance.

Table 5.2 Editors of surr and their appearance in the spaces used for discussing or editing the article

Article level	Remote editing		Automated editing			Discussion spaces		
en.surr	Wikimedia Commons	Huggle	Wikidata	Cluebot	Battybot	Wikimedia-l	Meta	RS/N
Utcursch	Utcursch							
Aprabhala	Aprabhala					Aprabhala	Aprabhala	
Fifelfoo	Fifelfoo							Fifelfoo
Sir Nicholas								Sir Nicholas
Battybot					Battybot			
MER-C								
Addbot			Addbot					
Cluebot				Cluebot				
Joseve		Joseve						

Those who attempted to ensure the fruitful travel of surr through Wikipedia's socio-technical system concentrated their efforts on the article itself, on the motivations for the project on meta and, to a lesser extent, on the Wikimedia-l mailing list. In terms of the article itself, edits by proponents of the fact took the form of adding text and other objects such as categories and images according to the norms of what was considered suitable to the project. The JPEG image was considered suitable enough for Hindi Wikipedia, for example, but when [Utcursch] translated the text of the surr article into English, he also translated the image into the norms of English Wikipedia where he believed that SVG images were considered the best format for images.

Another set of performative acts took the form of linking the fact to more powerful allies or traveling companions. In the context of surr, traveling companions included the article's sources (which, in turn, included both authors and subjects of sources, interviewers and interviewees) and the article's supporters. [Utcursch] regularly linked to the Oral Citations Project page in edit summaries and citations in order to bring in more powerful allies, but when those allies did not prove authoritative enough for those opposing the Oral Citations Project, he added the Jagran.com news source which remains unchallenged. References to other texts becomes an important way of stabilizing an article on Wikipedia. Without such references, articles become vulnerable to attack by virtue of their being published to a number of watchlists relating to lists of articles without citations, category links or incoming links. [Utcursch], himself, proved to be an authoritative traveling companion for the fact because his status as administrator enabled the articles that he published to be autopatrolled and so avoid scrutiny at the new pages level.

The Oral Citations Project page on the meta workspace contains reference to a number of authoritative allies from the Wikimedia Foundation and served as a way of framing

the problem that the oral citations were designed to solve. On the Wikimedia-l mailing list, proponents of the Oral Citations Project highlighted the disconnect between Wikipedia's goal of representing all human knowledge and the dearth of published material that reflected that knowledge. They praised the proposed solution in the form of oral citations for providing Wikipedia with the opportunity to be a different type of encyclopaedia by enabling subjects to have their knowledge represented directly. Oral citations, according to proponents, enabled Wikipedians to step into the role that had been (weakly) performed by academics in the past and to record the knowledge of ordinary people who they recognised as authorities on local culture.

Supporters of surr in its travels within Wikipedia were opposed by those who used the same tactics against them within the article, on the RS/N and on Wikimedia-l. The core problem that opponents seized on was that surr's only traveling companions came in the forms of sources that were disapproved of by the Wikipedians who tried to discredit them. In response to the actions of editors attempting to consolidate and strengthen the position of surr in the initial phase of the fact's travel, opposing editors attempted to label the article as unreliable and incomplete. They added warning tags to the article in order to cast doubt on the content of the article, they deleted text, thereby discrediting the fact and they removed the links to the oral citations hosted on Wikimedia Commons. They also threatened to delete the citations themselves by adding a request for copyright permissions in the format specified by policy and denied the authority of the Oral Citations Project page on meta by arguing that meta was irrelevant to the work of English Wikipedia.

Opponents to the oral citations labelled the authors and publishers of the oral citations sources as 'amateur' and the sources 'unreliable'. On the RS/N and Wikimedia-l, opponents contrasted oral citations to professional academic records in the form of

reports by ethnographers and historians who they positioned as the ideal sources. Opponents defined Wikipedians as passive collators of knowledge that exists elsewhere, rather as creative constructors of knowledge held by a variety of authoritative sources as defined by their environment. In this way, opponents destabilized the progress of surr by criticizing its traveling companions.

Both opponents and proponents of oral citations were driven by ideological strains and interests, but these strains and interests manifested themselves in different ways. For the proponents of oral citations, Wikipedia needed to change in order to accommodate knowledge from parts of the world that remained underserviced by the publishing market. For opponents of oral citations, allowing Wikipedia editors to also create sources had implications that were too far removed from the identity of the encyclopaedia.

[Fifelfoo] indicated in his interview with me that the sources themselves were badly constructed and were not reliable, but that he recognised the value of the project.

> I think it was a very very bad fit as a "reliable source" for the en.wikipedia project... (but) I am somewhat jealous that their project is closer to my understanding of how real knowledge is produced that en.wikipedia is.
>
> The "Encyclopaede" is like a Fordist factory of brutalisation and line work. And Wikipedia is like we've just taken over the factory for the workers... but the conveyer belts still discipline us. We accept the most heinous sources of knowledge on Wikipedia, as long as they're in a dead tree... or now represented at 1s and 0s. ([Fifelfoo] IRC interview, 9 February 2015)

This quote by [Fifelfoo], which was also used to set out this chapter, encapsulates a core idea: that Wikipedia's policies and practices are unmovable, and there is little that individuals can do using the tools that are available to build an encyclopaedia with. Although the Oral Citations Project recognises that individuals hold knowledge, rather

than experts or even the 1s and 0s that constitute published texts, [Fifelfoo] believed that this was how encyclopaedias are built and that this aspect was unmoveable.

5.3   The fact's terrain

Wikipedia's coded infrastructure plays an important role in defining the ideal environment for facts. Infrastructure, as constituted by policy and technological tools, determines who are 'good' traveling companions, what is 'suitable' terrain and which are 'characterful' or 'replicable' facts. Whereas policy (as outlined in chapter four) define at a foundational level what knowledge consists of and who is best placed to represent that knowledge, Wikipedia's coded environment served to reinforce particular interpretations of policy to advance certain ideologies above others. Tools are also used by editors to assert their ideological viewpoint on the article.

Wikipedia's coded environment benefits certain facts more than others. The surr example demonstrates that on Wikipedia, good traveling companions are those that are traceable, contained by numerous categories and linked to numerous other facts. When articles with no citations are published on English Wikipedia, for example, they are automatically flagged and added to a number of watch lists. In the 'new page feed', for example, where new articles are being evaluated, articles with no citations are flagged with the 'no citations' tag. Page patrollers can also search for articles to work through based on their lack of categories. Categories are themes that group a series of different articles together but some articles are not easily linked to other articles. Furthermore, articles with no links to them are classified as 'lonely' pages and can be deleted if they are not linked to other Wikipedia articles. Wikipedia's coded environment is therefore biased against facts that are not easily traceable online as well as facts that are not well

known by the majority of editors or that doesn't link easily to existing knowledge and can therefore not be easily discovered.

Tools and objects are also used by editors to assert their ideological viewpoint on the article. Inserting coded objects into the article served to evade increased scrutiny, for example. Adding the category, 'Traditional sports of India' and linking surr to the disambiguation page 'sur' served to remove it from particular watchlists that would involve further scrutiny according to the heuristics of different work groups in the surr case. Objects are also used to impose a particular lens on the article such as the application of warning tags to the surr article and the case when [Fifelfoo] admitted that he used copyright requirements as 'a stick to hit them (the Oral Citations Project participants) with', but that was not his main problem with the article. In order to be able to become active participants shaping facts within Wikipedia's environment, editors need understand how to stabilize or destabilize facts according to rules and norms. In particular, editors need to know how to use coded objects, either as weapons against or props to support the facts that they are either attempting to stabilize or destabilize.

The editors involved in successfully editing the surr article distinguished themselves by their ability to read, write and think with the language and tools of Wikipedia's socio-technical system. This language is expressed through the vocabulary and concepts used in the debates on the RS/N and on Wikimedia-l, it is expressed in the deployment of objects used to stabilize or destabilize the article, and through the use of gadgets and tools that provide some editors more powerful ways of thinking and acting with.

When editing the article, editors need to be able to recognise and translate traces made by other editors as well as by software agents in order to be able to conduct their own edits. The ability to read and write on Wikipedia is about being able to take account of how documents are circulated, interpreted, and evaluated. In the case of adding new articles, for example, it is important for editors to be able to read the traces that are created when the article moves through Wikipedia's evaluation processes in order for them to be able to contest actions taken against the article.

In the case of surr, traces made by some editors would have been difficult for newcomers to translate. The edit summary for MER-C's 7 March contribution, for example, is:

> 12:58, 7 March 2012 MER-C (talk | contribs)… (630 bytes) (-1,718)… (unsourced, fails WP:V, WP:RS)

Translating this summary, experienced users will be able to note that the time of the edit is in UTC and by the minute, thus facilitating discussion of edits that can take place minutes or seconds apart and providing a unique identifier for each version of the article that can be viewed by clicking on the date and time hyperlink. The editor's username is also hyperlinked, enabling editors to discuss the edit directly with them through their talk page. Providing a direct link to users' contributions ('contribs') also enables an assessment of an editor's point of view by those reviewing the edit. The number of bytes in the article is represented by the "630 bytes figure" while the "-1,718" refers to the fact that 1,718 bytes of information was removed by the editor. Both these figures are used in the evaluation of edits, with significant deletions of text figuring highly in algorithms that are used to recognise vandalism.

The edit summary, 'unsourced, fails WP:V, WP:RS' uses acronyms that are linked to Wikipedia's policies on verifiability and reliable sources. Each of these traces is the

compression of a number of rhetorical, normative, organizational and technical acts which enables the activities of editors to be aggregated, conceptualised and discussed. In this example, a single line in the edit history is a compact representation of a complex series of policies, practices and norms. Thus editors would need to look up these and other policies in order to be able to counter the edit. Furthermore, the assertiveness of the comment would make it difficult for editors to recognise that it could be alternatively interpreted in order to contest the edit.

Following discussions is also about knowing which are the most appropriate or sympathetic venues for either stabilizing or destabilizing the article. In the surr case, Prabhala and others discussed oral citations almost exclusively on Wikimedia-l but did not participate in discussions in the RS/N. The RS/N, however, turned out to be the critical forum in which decisions about oral citations were made and yet none of those who had created the article (namely [Utcursch] and [Aprabhala]) were present in the discussion. Additionally, those in the RS/N directed conversations to RS/N when they alerted editors to the discussion happening there on both the talk page and the Wikimedia-l discussion. Those in Wikimedia-l were not able to translate their feedback into the language of RS/N and thus failed to have a significant in countering the destabilisation tactics that were imposed on the article.

Writing on Wikipedia requires deploying objects and texts in a way that either evades or promotes visibility. An example of this in the surr article can be seen in Utcursch's ability to evade the new pages patrol process because of his status as an autopatrolled user. [Utcursch]'s autopatrol status meant that his contributions were not listed in the new pages patrol and were automatically listed as autopatrolled. He could therefore control the visibility of the article so that it was less subject to editorial oversight.

The use of powerful tools with which to edit also enables greater power over the writing process. Editors working in the areas designed to enable rapid evaluation of new content (including Huggle or the New Pages Feed) can potentially be much more effective in the sense that the scale of their edits can be significantly larger. The evaluation of new content in English Wikipedia is facilitated by a number of coded processes all taking place in distinct code/spaces, each of which enables different lenses for reading and writing. Tools have symbolic features (tagging an article for deletion indicates disagreement, opposition, challenge to either the idea or the author/s of an article, for example) but also materially affect what a user is able and encouraged to do or say within them.

For Huggle users, new edits to surr would have been reflected in the interface with particular characteristics deemed relevant for the identification of vandalism highlighted by the interface. Such characteristics or heuristics include the size of the edit in bytes and a "user badness score" indicating the probability that the highlighted user was making a bad edit. Similar characteristics are highlighted when the article appears in the New Pages Feed. Page patrollers are able to filter new pages according to a number of set variables. They can show unreviewed pages, reviewed pages, pages nominated for deletion or new redirects.[35] They can also specify articles created by specific editors by inputting their username. The evaluation of content in both Huggle and the New Pages Feed is designed in a way that prioritises rapid decision-making on

---

[35] Redirects are pages that contain no content but send the reader to another page.

multiple articles[36] thus enabling editors working in these areas to become more effective and more prolific actors than those not using these tools.

The vocabulary used on Wikipedia is highly technical. Adding citations to an article in a way that the software will recognise requires the use of citations templates, for example. If citations are not added using citation templates, then they won't be recognised by the software and therefore the article could be automatically flagged for greater oversight which could put it at risk of destabilization.

Knowing how to produce traces, to employ gadgets that improve the speed and efficiency of the editing process, to read the traces produced by others, to understand which discussion spaces are most relevant to controlling the stability of the article are all necessary skills in order to become an effective contributor to Wikipedia. In turn, evaluations of Wikipedians by other Wikipedians is taking place regularly in code/spaces such as the New Pages Feed and Huggle, where an editor's edit count, and the content of their contributions is seen as relevant to assessing their credibility. Expertise is therefore a self-fulfilling prophecy; the more editors can effectively contribute, the more they will be recognised as authoritative and have their edits prevail.

5.4 Conclusion

The Oral Citations Project presents an enormous challenge, not only to Wikipedia, but to Western ideals about who has the authority to represent facts that are coming to be

---

[36] On 9 November 2014, 7,873 pages were still marked as unreviewed and the oldest had been produced over a year ago.

relied upon. Advocates of the project chose players of surr, rather than historians or anthropologists to be interviewed for the surr article, effectively upending Wikipedia's notions of expertise as located within institutions, particularly academia. The project advocates were effectively suggesting that knowledge could be gained by the embodied experience of phenomena rather than through institutional certification.

This idea represented a threat, not only to the series of articles using oral citations, but to Wikipedia's conceptions about its own role in the system of expertise more broadly. By choosing to interview individuals in the community who had first-hand knowledge of the game, the editors representing surr in the original Hindi had effectively transformed those community members into experts. Similarly, the editors who conducted the interviews also displayed expertise because of their knowledge of the phenomenon existed and knowledge of how to identify and contact the experts in the villages.

The response by some English Wikipedia editors was to reaffirm the expertise of academics and institutions, thereby denigrating the embodied experience of the villagers. When the article entered English Wikipedia's socio-technical system, the voices of those who experience the game were no longer available since the link to the audio file had been removed. Instead of the community members being recognised as experts, they were relegated to being the subjects of the newspaper article about surr that constitutes the only linked citation within the article when this study was conducted. The Wikipedia editors who were able to deploy their allies, cut off access to the allies of their opponents, deploy coded objects and choose the most effective forums for their arguments were most successful. Their success lay in their ability to define the roles and identities of Wikipedians and the experts that they rely on.

As Jasanoff (2004) writes, power is continually reinscribing itself in institutions, practices, discourses, claims and products of science and technology (p. 36). The example of surr demonstrates how the power to represent facts is lodged in discourses concerning the sources of reliable knowledge, representations that determine who the expert and subject are, and identities that authorise certain knowledge holders as legitimate and others as illegitimate. In this case, editors of surr employed discourses in order to limit what could be said about what Wikipedia includes. For these editors, the most important factor in determining inclusion into English Wikipedia was whether the source was verifiable as defined by Wikipedia policy. Editors successfully determined that discussions about market failure in the publishing arena were irrelevant to such discussions.

Power was also lodged in representations of surr, where there were disagreements about who surr's traveling companions could legitimately be, who was designated the role of expert supporter of the fact and who was designated the role of passive subject. Opponents of the oral citations designated surr's spokespersons (represented in the oral citations interviews) as illegitimate, and did not challenge the newspaper interview, where the journalist was the designated spokesperson and considered legitimate.

Finally, power was lodged in the identities of those representing surr. In discussions, opponents were able to label Wikipedians as unreliable sources or publishers of information, and to relegate editors to the role of recorders of sources published elsewhere. It became clear in the strategic power play engaged with by editors, however, that a considerable amount of expertise was required in order to stabilize or destabilize the article and thus participate fruitfully in the travel of the fact. Although Wikipedia editors have to outwardly deny their own knowledge and its reflection in the

article, it is clear that editors are, in fact, demonstrating the active application of their knowledge to the article in either urging it forward to destabilizing and thus preventing its travel.

The next chapter describes the progress of a fact that moved far further within Wikipedia's infrastructure. This chapter follows the progress of the 2011 Egyptian Revolution article over time, taking note of the actors, objects and processes that attached themselves to the facts as they travelled over time. The chapter analyses how well the fact travelled, which traveling companions were considered authoritative and what the implications of this are for the reconfiguration of expertise in the digital environment.

Chapter 6: The rapid and fruitful travel of the Egyptian Revolution of 2011 facts

> I didn't make these rules, it's how Wikipedia operates.
>
> ([.68], Egyptian Revolution of 2011 talk page [37], 3 February 2011)

Up until this point, the discussion has focused on a fact that struggled to travel far from its site of first publication on Hindi Wikipedia. In this chapter we turn our attention to a fact that spread with great velocity, integrity and fruitfulness from its first publishing on English Wikipedia and ask why it enjoyed this trajectory.[38] By investigating how the fact's traveling companions, boundaries and terrain were perceived and constructed as it travelled through Wikipedia, we also ask what the impact of discourses and infrastructural biases are on the ways in which authority and expertise are being reconfigured.

The Egyptian Revolution of 2011 facts have travelled far and wide within Wikipedia over the past four years. As of the time of this study, the Egyptian Revolution of 2011 article exists in 58 different language versions of Wikipedia. There are 57 redirects to the English Wikipedia article (indicating alternative titles for the article) and almost 2,000 other articles and pages within Wikipedia link to the article.[39] Almost 2,000

---

[37] The archive of talk page discussions for the Egyptian Revolution of 2011 article can be found at https://en.wikipedia.org/wiki/Talk:Egyptian_Revolution_of_2011.

[38] The article was first published on English Wikipedia at 13:26 UTC on the 25th of January 2011; the Arabic version was published about 3.5 hours later at 16:55 UTC and an Egyptian Arabic version was published three days later on the 28th of January.

[39] See https://en.wikipedia.org/wiki/Special:WhatLinksHere/Egyptian_Revolution_of_2011.

editors have made just over 8,000 edits to the Egyptian Revolution of 2011 English Wikipedia article as of 5 May 2015[40] and at least eight new articles were started in order to extend the description of facts introduced in the article. These features make the Egyptian Revolution of 2011 article useful for studying why certain facts are able to travel fruitfully and with integrity within Wikipedia.

There are both similarities and differences to the evolution of facts on Wikipedia that are related to breaking news and articles about phenomena and events constructed after their appearance. Keegan, Gergle & Contractor (2013) performed network analysis on both breaking and non-breaking articles on English Wikipedia, defining the breaking status of articles in terms of the 'temporal proximity between the article's creation and the date of the incident itself' (p. 7). Similarities included that both types of collaborations enjoyed similar numbers of contributors, similar contributors and similar distributions of activity but only in the long term. Different from non-breaking articles, breaking articles were found to exhibit a rapid role specialization among editors who have worked together in the past in the initial stages of the event.

The authors note that more research is needed to understand how this happens since 'the implicit or explicit processes by which these roles are negotiated and the social or cognitive channels through which they are shared remain unclear' (p. 25). Without managerial command structures, editors seem to rely on other mechanisms to coordinate their activities, but the mechanisms by which this happens have yet to be discovered.

---

[40] See https://en.wikipedia.org/w/index.php?title=Egyptian_Revolution_of_2011&action=info.

This chapter sheds some light on the ways in which coordination mechanisms are established early on in the construction of an article. In the first stages of the Egyptian Revolution of 2011 article construction, particular sets of editors established themselves as coordinators and leaders of the article. The authority of these editors to coordinate the article originated in their experience of the topic (for example, those who were among the protesters in Egypt) or their experience of Wikipedia editing (as demonstrated by the deployment of policy and other editorial objects). As leaders of the article and the facts contained with it, experienced editors were able to stabilize the article according to their own standards. These standards weren't entirely constructed by leading authors but were, instead, guided by the policies and norms discussed in chapter four.

Editors who took on this leadership role were able to deploy policies and other editorial objects in silencing their critics or simply in representing facts according to their own ideals. Significant to such representational standards was the ability to establish which sources were considered good traveling companions within the context of the article. Sources that did not fit the ideal were removed over time and replaced by sources that editors felt were more reliable or that they believed were more reliable to their readers.

The Egyptian Revolution of 2011 facts travelled well within Wikipedia's socio-technical structure because their terrain was established by the existence of numerous, highly accessible sources that fit (loosely) within Wikipedia's policy definitions of reliable sources and fit the practices by which editors discovered sources during events. Furthermore, the character of the fact itself (that it fitted a narrative that was approved of by the West) also helped to catalyse the rapid travel of the fact beyond its first establishment as an article on English Wikipedia.

This travel was not obvious, nor was it predictable. Wikipedia's policy against writing about events too soon after they have occurred has been used in the past to prevent the establishment of a fact through article creation, but the factors mentioned above proved too strong to prevent this fact from moving rapidly through Wikipedia. This chapter follows the progress of the Egyptian Revolution facts as they moved through Wikipedia's infrastructure. The fact's traveling companions, as well as its terrain and character are analysed, with the final section asking why the fact was able to enjoy such wide travel within Wikipedia, and the implications of this for the ways in which the power to represent are being lodged in particular parts of the network of sources that accompany facts on Wikipedia.

6.1   An account of the fact's travels

The primary fact under analysis is embedded in the article title, currently the 'Egyptian Revolution of 2011', but previously titled '2011 Egyptian protests' and also known by a number of alternative titles that redirect to this article. The primary home of this fact, however, and one which was hotly debated by editors is at https://en.wikipedia.org/wiki/Egyptian_Revolution_of_2011 that designates events that took place from January 25, 2011 in Egypt as revolutionary. This primary fact is supported by numerous other facts that make up the content of the article; facts also occur in categories and images, captions and infoboxes as can be seen in figure 6.1.

Figure 6.1 Facts and traveling companions of the Egyptian Revolution of 2011 article

Source:
https://en.wikipedia.org/w/index.php?title=Egyptian_Revolution_of_2011&oldid=674520833

The subjects of the facts are the people of Egypt, some of whom gathered in Tahrir Square in Cairo and other points throughout the country during events in 2011. The voices of these subjects have been archived in numerous social media archives since social media sources including Facebook and Twitter were central platforms for the spreading of messages by both opponents and proponents of then President Hosni Mubarak. Events were also covered by local and foreign journalists who streamed into Egypt from January 2011 in order to witness the protests. Google Scholar also lists thousands of scholarly articles written about the 'Egyptian Revolution of 2011' in the past four years.

All of these representations, from social media, journalistic and academic sources, were potential traveling companions for facts on Wikipedia, but certain types of sources were

203

prioritized by editors. Such prioritization was determined by the fact's terrain and its character, with its terrain constituted by the practices that Wikipedia editors engaged in in editing the article, as well as the availability of sources and citations and the particularities of the fact that made it glide easily through Wikipedia's socio-technical system from January 2011 onwards.

6.1.1   The first edits

On the 24th of January, Egyptian activist and Wikipedia editor, [The Egyptian Liberal] prepared an article for English Wikipedia entitled '2011 Egyptian Protests' for posting the next day. Although local media in Egypt as well as blogs and social media were publishing news stories about calls for protest on National Police Day in Egypt the next day, there were few stories in the international media covering planned events. In fact, in the week prior to the first mass protests, both Time Magazine and the BBC published articles explaining why Egypt would *not* follow protests in Tunisia that had recently unseated the government of President Zine El Abidine Ben Ali (Hauslohner, 2011; Leyne, 2011).

In my interview with him, [The Egyptian Liberal], who was a university student at the time, noted that believed that the protests the next day would be significant enough for an article on Wikipedia.

> **[The Egyptian Liberal]:** I thought the thing was going to be big… before the revolution became a revolution. Two days before I thought: this is going to be big.
>
> **[HF]:** Why did you think that?
>
> **[The Egyptian Liberal]:** The frustration in the street. And especially what happened in Tunisia. There were a few self-immolations in Egypt that weren't covered in the media and by bloggers as much as Bouazizi in Tunisia. When people set themselves on fire you know something is seriously wrong with the society you live in. (Skype interview, 14 October 2012)

The first draft would not have triggered any warnings in the new article feed because it contained the three elements (citations, categories and links to other Wikipedia articles) necessary to evade warning flags. The citation, however, was to an Agence France-Presse (AFP) news report that described the events in the future, that is, how Egypt was bracing for a day of nationwide anti-government protests the next day (AFP, 2011) rather than a report of what happened on the day of the protests themselves. The article also linked to other Wikipedia articles through the addition of categories including 'Riots in Egypt' when there was no evidence of riots at the time, and the image was a cartoon by political cartoonist, Carlos Latuff (see figure 6.2) when the space would ideally have been a filled by a photograph of the protests.

Figure 6.2 Infobox from the first version of the Egyptian Revolution of 2011 (then 'Protests') article

Notes: The cartoon's caption read: 'Khaled Mohamed Saeed holding up a tiny, flailing, stone-faced Hosni Mubarak'. Khaled Mohamed Saeed was a young Egyptian man who was beaten to death reportedly by Egyptian security forces and the subject of the Facebook group 'We are all Khaled Saeed' moderated by Google employee, Wael Ghonim. Ghonim used the Facebook group (Ghonim, 2012) to call for a revolution on January 25 2011, when people took to the streets to protest against the government of then President Hosni Mubarak.

Source:
https://en.wikipedia.org/w/index.php?title=Egyptian_Revolution_of_2011&oldid=409962868

One of Wikipedia's foundational policies mentioned in chapter four, 'What Wikipedia is Not',[41] states that 'Wikipedia is not a crystal ball' so that predictions or speculations about events should not be included. Even though the editorial elements published in the first version of the article might have broken Wikipedia policies, the article contained the relevant coded objects necessary in order to avoid immediate warning flags. Three prominent cleanup tags were, however, added to the article when it was first published. These included the 'new unreviewed article' tag, the 'expert in politics needed' tag and the 'current event' tag, warning readers that information on the page may change rapidly as events progress. These three lines of code that constituted the cleanup tags initiated a complex distribution of tasks to different groups of users located in work groups throughout the site: page patrollers, subject experts and those interested in current events. Coordination among established Wikipedians who were notified of the article through their membership of these work groups had thus begun.
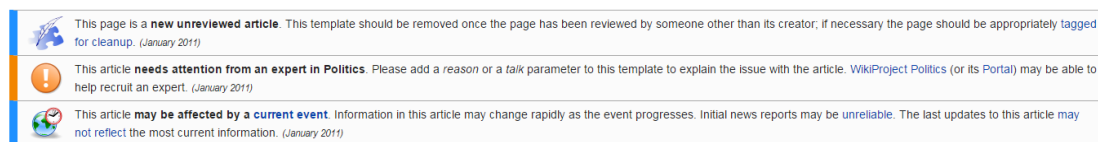
Figure 6.3 Screenshot of the three cleanup tags automatically appended to the head of the Egyptian Revolution of 2011 article when it was published on 25 January 2011

Source:
https://en.wikipedia.org/w/index.php?title=Egyptian_Revolution_of_2011&oldid=409962868

The first tag was automatically appended to the article when it was created using the 'Article Wizard', a series of forms that take new editors through the process of creating an article.[42] This was the first (and perhaps the most significant) challenge faced by the fact. As soon

---

41 See https://en.wikipedia.org/wiki/Wikipedia:What_Wikipedia_is_not.

42 See https://en.wikipedia.org/wiki/Wikipedia:Article_wizard.

as an article is published on Wikipedia, it is added to a list of unreviewed articles watched by page patrollers who then review the article and determine whether it qualifies for speedy deletion, whether it should be nominated for deletion (involving a more lengthy, consultative process) or whether it should be allowed to continue.

The second automatically appended cleanup tag (that help was needed from an expert in politics) alerted users in the WikiProject Politics group to the existence of the article when it appeared in a list of other articles seeking assistance from WikiProject participants. The third cleanup tag ('current related') alerted users to the rapidly developing nature of information on the page and added the article to the category of pages entitled 'current events'.

Despite the existence of these warning tags, the article endured. When I asked in an interview why he thought that the article wasn't deleted, [The Egyptian Liberal] replied that he had warned other editors about what was happening and that the article would be coming.

> **[HF]:** And the article was never put up for deletion or anything like that, right?
>
> **[The Egyptian Liberal]:** No, because I already… I made few editors aware of the article and I had enough sources there, so nobody can actually claim that there's not enough sources for a significant event, so… and I had enough people who did post. Their only responsibilities were to take pictures or upload pictures from different Flickr accounts to back the claim that, well, people are protesting.
>
> (Skype interview, 14 October 2012)

Whether or not this was the reason that the article remained unchallenged is unclear, but what is clear is that [The Egyptian Liberal] had access to a communication network outside of Wikipedia's visible procedures that enabled him to be successful in his editing practice. As the day progressed, [The Egyptian Liberal] continued to fill out the background section of the article, listing individuals who had set themselves on fire in protest against worsening economic conditions in Egypt, and adding to the growing list

of references. About two hours later another editor joined him, removing the new article template and political expert needed tag, improving the grammar, adding links to other Wikipedia articles and listing categories. The actions of these two editors at the beginning stages of its development were crucial in solidifying the long-term future of the article.

6.1.2   Controversies

The next two weeks constituted the most popular editing period for the article. This was helped by the suspense attached to news reports at the time, as people waited to hear whether then President Hosni Mubarak would resign in compliance with the protesters demands or whether there would be further retaliation by the State. Editors regularly updated one another about the events as they were reported by local and international media organisations on the talk page of the article. Although the majority of discussions indicated consensus, there were a few occasions in which editors disagreed with one another, resulting in lengthy, often passionate debates.

Three of these debates are highlighted below. They include a debate about whether a table used in the article to account for deaths as a result of protests constituted original research according to Wikipedia policy, a debate about the editorial distance that needed to be established between Wikipedia and its sources, and a debate concerning how the article's main title should be decided. Each of these debates symbolises disagreements about Wikipedia's identity as an encyclopaedia as well as the role of editors as either passive aggregators of facts determined by others, or as active participants in the construction of the narrative about events in Egypt in 2011.

Keeping track of death counts in the article was highly contentious because no single external source was maintaining a comprehensive list of the accumulating deaths. Numbers needed to be stitched together by editors but this resulted in inconsistencies when different editors reflected numbers from different sources or were adding numbers incorrectly. The debate came to a head when anonymous user, [.68] complained that numbers of dead in the article contradicted numbers in the table (see figure 6.4) and that the table should be removed. When [.68] did not feel the issue was being taken seriously enough, s/he deployed a policy object into the discussion and argued that editors were breaking the rules of original research by producing a table that was not reflected in any independent source.

> **[Wipsenade]:** I'll call it 'Unconfirmed total' then, to be safe.
>
> **[Physics is all gnomes]:** sounds like a plan :) -
>
> **[The Egyptian Liberal]:** Agree per Wipsenade, Physics is all gnomes and Lihaas
>
> **[.68]:** It's not a matter of local consensus, it's the "three pillars of Wikipedia". You can't create any original content not backed by just what the reliable sources are saying, not in such cases as creating history.
>
> And yet it would be still OR,[43] meaning it's unacceptable here. You can post "[figure] people reportedly died in Cairo so far, according to unconfirmed reports" alright, but only when a reliable source says *just this*. But you can't add and sum various unrelated reports using *sloppy math* (that's a quote), because it's constitutes original reasearch, and this is a big no-no. I didn't make these rules, it's how Wikipedia operates.
>
> (Egyptian Revolution of 2011 talk page, 3 February 2011)

The majority of editors acquiesced and the table was removed with [Wipsenade] writing:

> OK, it's out untill it's confirmed in time by the BBC, UN, EuroNews, CNN, NHK, Ruiters, ect. ([Wipsenade], Egyptian Revolution of 2011 talk page, 3 February 2011)

---

[43] Original research.

Major flash point	Sourced death toll by 29 January 2011	Sources
Alexandria	29	[84][98]
Suez	28	[84][98]
Beni Suef	17	[97]
Thebes	1	[138]
Atfih	1	[139]
Cairo	15	[84]
The rest of North Sinai Region	2	[61]
Sheikh Zoweid	1	[60]
Aswan	0	[84]
Luxor	0	[98]
Rafah	3	[98][98]
Giza	0	[84][98]
Sharm El-Sheikh	0	[98]
Hurghada	0	[98]
Mansoura	2	[140]
Other places hit by protests (if any)	14	[61][84][98]
Total	136	

Figure 6.4 Death count table from the 31 January 2011 version of the Egyptian Revolution of 2011 Wikipedia article

Source:
https://en.wikipedia.org/w/index.php?title=Egyptian_Revolution_of_2011&oldid=411157336

A second key debate among editors occurred on the 10th of February, a day before then-President Hosni Mubarak resigned from office. [Ocaasi] announced on the talk page that he had added a link to Al Jazeera's YouTube channel with live reporting from Tahrir Square. [Ocaasi] had placed the link in the 'Further reading' section of the article but it had been removed by another editor. [Ocaasi] asked why this happened on the talk page, explaining that he would normally have discussed his addition there first, but since the event was happening 'right now', he was 'going to put it (the link) back' ([Ocaasi], talk page, 10 February 2011).

The debate that ensued involved editors discussing the extent to which Wikipedia should be a passive 'mirror' of media organisations such as Al Jazeera or whether they should be recording events only after they have been analyzed by reliable sources.

[Ocaasi]: Maybe the recentism is part of the issue, but I don't see this as compromising our encyclopedic stance. For one, this entire article is recent, and secondly, external links have a much looser approach to content which is useful but not directly encyclopedic. Thoughts?

[Cptnono]: You added a link to the Chanel of Al Jazeera. It is interesting but will not be related to this article as soon as their main video is something else. We are not here to spread breaking news... Although this is interesting stuff, it turns the page from an encyclopedic article to a mirror for news sites. If someone is on (the) internet (right) now, they can see similar live coverage at cnn.com or multiple other places.

(Egyptian Revolution of 2011 talk page,10 February 2011)

[Ocaasi] responded that the Al Jazeera live video link was a unique resource for a momentous occasion and so rules should be relaxed but [Cptnono] threatened to report him to administrators if he broke the three revert rule:[44]

> **[Ocaasi]:** 1. It's is a unique resource right now. 2. It's not being emphasized, just included 3. IAR. The speech is live, and this is coming right from Tahrir square. I'm putting it back, but will be happy to debate the policy subtleties tomorrow when the event has passed. I will also then remove the link.
>
> Some outside opinions would be welcome here, since I don't want to break 3RR as this continues, but I'm willing to make an exception for an event as momentous as this.
>
> **[Cptnono]:** And I am willing to report you for edit warring. If you want people to see the video post it on a blog. This is not a blog.
>
> (Egyptian Revolution of 2011 talk page,10 February 2011)

Over the next two days the argument continued, with [Cptnono] requesting advice at the External Links Noticeboard, and [Ocaasi] requesting advice on Jimmy Wales' talk page, until three editors brought external link content guidelines into the discussion in order to oppose the use of the link. After [Cptnono] stated that he was 'removing it finally' on

---

[44] The three revert rule on Wikipedia prevents editors from reverting one another's edits more than three times in a 24 hour period and is an attempt to prevent spiraling edit wars where editors continuously revert one another's edits in a process that continually destabilizes an article.

211

the 12th of February 2011, there are no further comments. [Cptnono] disallowed the live link because, to his mind, Wikipedia's hosting of live sources steered too far from Wikipedia's definition as an encyclopaedia. [Cptnono] was thus able to establish what s/he believed to be the appropriate distance between Wikipedia and its sources using allusions to policy and by threatening administrative sanctions.

The final example of a controversy that occurred within the talk page of the article occurred when then-President Hosni Mubarak resigned on the 12th of February 2011. Editors had been calling for a name change to the article before this, with some editors opposing a change to 'revolution' until Mubarak had resigned. Scores of editors attempted to change the name of the article from '2011 Egyptian protests' to '2011 Egyptian Revolution' minutes after the announcement had been made by Mubarak that he was stepping down. One editor, [J4V4] initiated a poll on the talk page in which editors were asked to weigh in with reasoned arguments about whether the name of the article should change to 'revolution'. At the heart of the argument was a question about whether a revolution had, indeed, occurred in Egypt and how to assess what was the predominant title used by reliable sources in describing the events.

In parallel to the discussion, a crowd took over, making the name change repeatedly and being repeatedly reverted. Editors continued to add their voices to the discussion over the next 24 hours but the move was done just three hours after the proposal was made even though the decision on the talk page was recorded 24 hours later. This example demonstrates how the sheer numbers of actors demanding particular edits can overwhelm the deliberative process and that, what seems to result from a deliberate process, can have already been decided through the speech actions of editors working directly on the article.

The majority of the 2,000 editors, about 1,500 citations and almost 8,000 edits that constituted a major portion of the actor network of the Egyptian Revolution of 2011 article were active in the first few weeks of the events (see figure 6.5). As demonstrated in the controversies above, also relevant to the travel of facts were policy objects and disciplinary procedures that were used to empower certain actors to impose their representations of facts on others. These actors, objects and their attendant processes and practices constituted the facts' traveling companions and its terrain.

Number of edits to the Egyptian Revolution of 2011 Wikipedia article in 2011
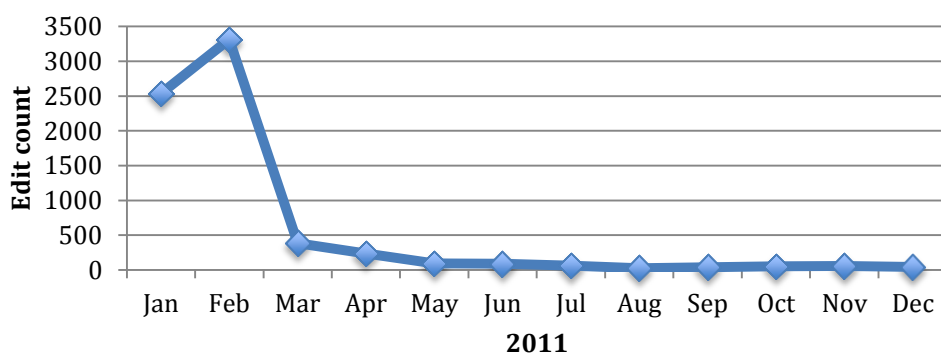
Figure 6.5 Numbers of edits to the Egyptian Revolution of 2011 article on English Wikipedia in 2011

Source: author

6.2   Traveling companions

6.2.1   Editors

At the core of the actor network catalysed by the first version of the Egyptian Revolution of 2011 article are the thousands of editors and hundreds of citations to sources that populated the article over time. Editors added phrases, paragraphs, tables, infoboxes,

213

images, categories, and headings to the article as they encountered and engaged with sources of information outside of Wikipedia. By applying what they knew about Wikipedia policy, norms, and technical architecture, as well as their individual perceptions of the credibility of sources and the authority of those sources to speak on subjects specific to the article, editors made both individual and collective decisions about which sources to select, which to challenge, debate or delete, and which to allow.

Almost 2,000 editors have made edits to the Egyptian Revolution of 2011 English Wikipedia article as of 5 May 2015, with the majority of edits taking place in the first two months of 2011. 25% of the article's editors edited anonymously under their IP address. The majority of editors originated from the United States, but there were also anonymous edits from Egypt, Poland, the United Kingdom and Canada. Fewer than 1% of the editors of the page indicated that they were female. The article is quite typical for Wikipedia articles relating to breaking news topics. The majority of articles in this category are edited heavily during the first stages while events are highly visible in the media (Keegan et al., 2013). Anonymous editing for the article is close to the site's average at 29% (stats.wikimedia.org) but the proportion of female editors is low (at 1%) in relation to the site average at between 10 and 16% as indicated in chapter one. [45]

Editors involved in the construction of the Egyptian Revolution of 2011 article were motivated by a variety of factors. According to one of the most prolific editors, [Ocaasi], 'there's a wild, addictive rush to working on an article when tens of thousands of people are coming to read about it' ([Ocaasi] Skype interview, 3 April 2014). For other editors, such as [The Egyptian Liberal], editing the article was a way for him to be involved in

---

[45] The 1% figure only accounts for those who indicate their gender on their user page and so the figure could be higher.

writing the history of his country and of paying tribute to those who perished in the protests. In a statement on the talk page of the article, [The Egyptian Liberal] concluded that his personal biases were driving him towards editing neutrally rather than in a biased manner.

> I think I owe the people who are protesting and those who have passed away to tell their story from a NPOV (neutral point of view). I am not nor can I be in Egypt due my university so that my way of standing up to Mubarak and tell him *I shall not be silenced and I shall tell the world what happened*. ([The Egyptian Liberal],[46] Egyptian Revolution of 2011 talk page, 10 February 2011)

The need to be involved in the historic retelling of events (in additional to other personal and professional interests) motivated hundreds of editors to participate in the editing of the page, with a few editors working night and day to manage the overall direction of the article as events evolved.

Although almost 2,000 editors offered some kind of contribution to the article, a handful of editors made the majority of edits. Figure 6.6 below illustrates edit counts of the top 10 editors of the article who contributed almost 70% of the edits to the article. A core group of editors established themselves as an ad-hoc editorial committee during the first few months of the article's development, with a few stable members including the original author of the article, [The Egyptian Liberal], [[Lihaas]] and [Ocaasi].

---

[46] In my interview with him, [The Egyptian Liberal] said that he actually *was* in Egypt at the time. He may have made this claim on the public talk page for security reasons.

10 most prolific editors of the Egyptian Revolution of 2011 Wikipedia article
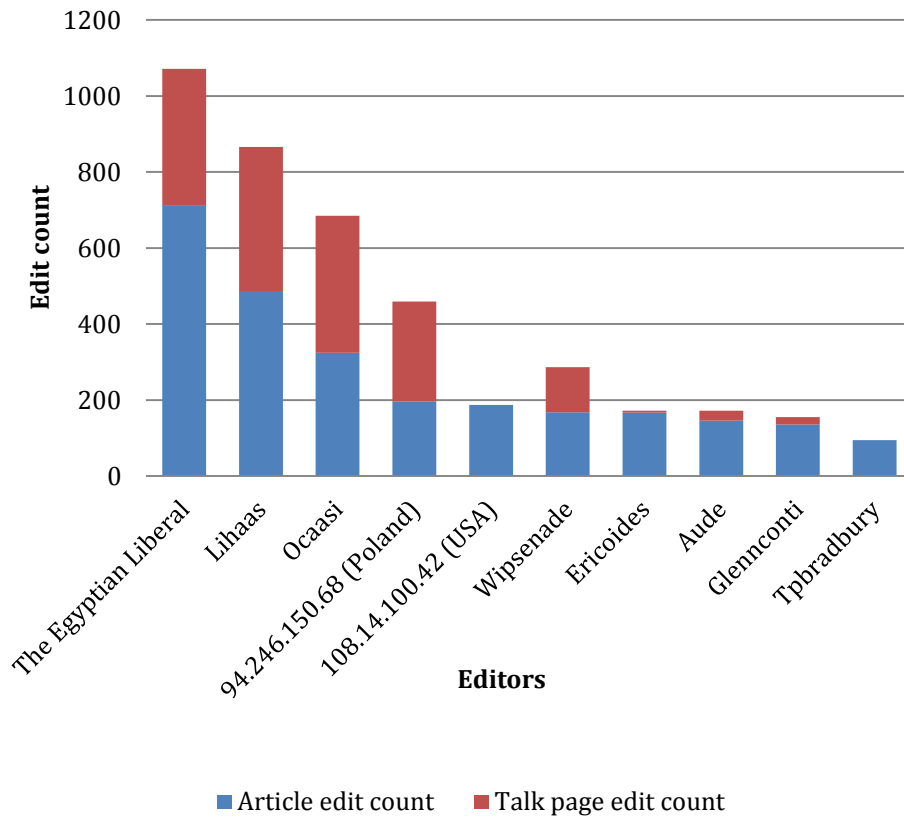
Figure 6.6 10 most prolific editors of the article and talk page of the Egyptian Revolution of 2011 article (as of 5 May 2015)

Source: author

Figure 6.6 also illustrates how contributions to talk pages and to the main article space were relatively similar for a few editors such as [Lihaas] and [Ocaasi], but that the majority of the most prolific editors of the article did not participate actively in discussions on the talk page. [Aude], for example, contributed 146 edits to the article but only 26 talk page edits. A few individuals including [The Egyptian Liberal], [Ocaasi], [Lihaas] and anonymous editor whose IP address resolves to Poland, [94.246.150.68],[47]

---

[47] Named [.68] from here onwards.

dominated editing of both the article and the talk pages, becoming powerful coordinators of the overall narrative over time. Other editors such as anonymous United States-based editor, [108.14.100.42], preferred to avoid discussions and debate on the talk page and concentrated, instead, on editing the article.

[The Egyptian Liberal], the user who started the article, was the most prolific editor by a significant majority. Along with [Ocaasi] and [Lihaas], [The Egyptian Liberal] became the de facto leader of the page, making regular announcements about the state of the article, responding to requests for edits on the talk pages, and intervening in conflicts between editors. In my interview with him, [The Egyptian Liberal] stated that he shared his Wikipedia user account with three other people in different countries during events in Egypt.

> During the revolution it was not only me editing from the page. I had different people who didn't have an account on Wikipedia who wanted to edit Wikipedia especially at the time of the revolution they really wanted to help. So we were three different people with different writing styles... The two that were helping most were not Egyptian. ([The Egyptian Liberal] Skype interview, 14 October 2012)

[The Egyptian Liberal] demonstrated a keen awareness of how the traces that he made on Wikipedia would be visible to people who may use those representations against him depending on which way events unfolded in Egypt. This enabled [The Egyptian Liberal] to evade at least some of the potential surveillance that his traces would be subject to by sharing his IP address.

In the hours, days and weeks following the creation of the article, editors (some of whom had worked together on other articles relating to Middle East politics before this) quickly settled into a routine, falling into roles based on their strengths and expertise in relation to one another and to the needs of the article. Many of the editors were avidly watching the live feed of events on Al Jazeera television or online via Al Jazeera's

YouTube channel as they edited, and editors would edit in informal shifts in order to cover events as they progressed.

[Ocaasi] and [The Egyptian Liberal] formed a partnership in leading work on the page. [Ocaasi] is a native English speaker based in the United States and had seriously started editing Wikipedia a few years before. [The Egyptian Liberal] is not a native English speaker but, as the only significant Egyptian contributor to the article on English Wikipedia and with personal connections to some of the protest groups involved in demonstrations, was able to play a key role in highlighting who was notable in Egypt's political environment and to verify information that required contextual knowledge.

> We developed a highly symbiotic relationship. ([The Egyptian Liberal]) knew who was an important in Egypt… he had the connections while I had the ability to research and write particularly well in English. ([Ocaasi] Skype interview, 3 April 2014)

[Aude] was a seasoned American Wikipedian who had lived in Egypt before and had some knowledge of the political context there. She maintained a list of reliable sources on her talk page, chimed in on debates and coordinated attempts to get Al Jazeera to license their photographs under Creative Commons, when the editors were struggling to find openly licensed images. [SilverSerenC] was an American university student at the time who worked on finding sources, verifying information and contributing to debates on the talk page. [Lihaas] contributed significantly to the talk pages and was a frequent editor of the article.

3 most prolific bot editors of the Egyptian Revolution of 2011 Wikipedia article
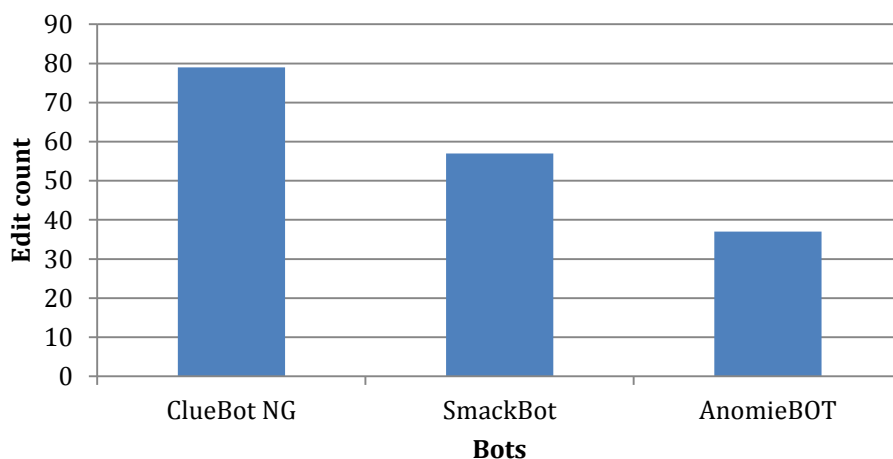
Figure 6.7 The three most prolific bot editors of the Egyptian Revolution of 2011 English Wikipedia article

Source: author

3.3% of all edits to the article as of 5 May 2015 were by bots (see figure 6.7). Three of the most active bots that edited the article (see figure 6.7 above) included [[ClueBot NG]] (a vandalism detection bot), [[SmackBot]] which corrects ISBN representations (among other editorial tasks), and [[AnomieBOT]] which corrects pages with incorrect reference formatting, among other tasks. The role of bots was to automatically correct elements of the article according to a set of predetermined rules. Like the bots operating on the surr article, these bots reordered content in ways that impacted the readers' and editors' perceptions and activities in relation to the facts on display. [[ClueBot NG]], for example, developed an algorithmic portrait of what vandalistic content looks like and applied numerous rules (as well as its experience of detecting vandalism before this event) in order to determine what content was valid.

6.2.2 Citations

Editors use citations during breaking news events as a way to back up claims being made, and as a trail for users to follow in order to verify whether the sources were being accurately reflected. The discovery, triangulation, selection and summarising of sources is central to the practice of article editing and sourcing practice plays an important role in determining which sources are more prominently displayed than others.

Discovery is a process in which editors open themselves to streams of facts about an event that are originating from particular sources or search for sources that support particular facts. Discovery is the first step in a process that begins with the editor's intention to find novel facts to add to the article, or as a way of altering facts based on what they have discovered. Editors always come with existing knowledge or preconceptions about the subject of the facts under analysis so that they sometimes appear to be discovering sources and facts, but in many cases are searching for evidence from sources that fit their conceptions of what the facts are.

One editor who I spoke to in my earlier research said that the citation reflected in the article by an editor is not always the source of that information. Many editors will discover information from local sources, for example, but then replace those sources with sources that are more commonly accepted on Wikipedia such as those of international publications such as the New York Times and the BBC (Ford, 2011).

At the time of data collection (14 July 2014), the 'Egyptian Revolution of 2011' article on English Wikipedia contained 1,565 citations. 97% of the citations are represented by a URL, but the remaining 3% are mostly books whose full text is not available online. There is a long tail of sources represented in the article, with more than half of the

220

citations representing only a single domain (see figure 6.8). Previous quantitative research that I conducted with colleagues indicates that sources on English Wikipedia follow the same long tail structure (see Ford et al., 2011).

Domains for citations added to the Egyptian Revolution of 2011 Wikipedia article
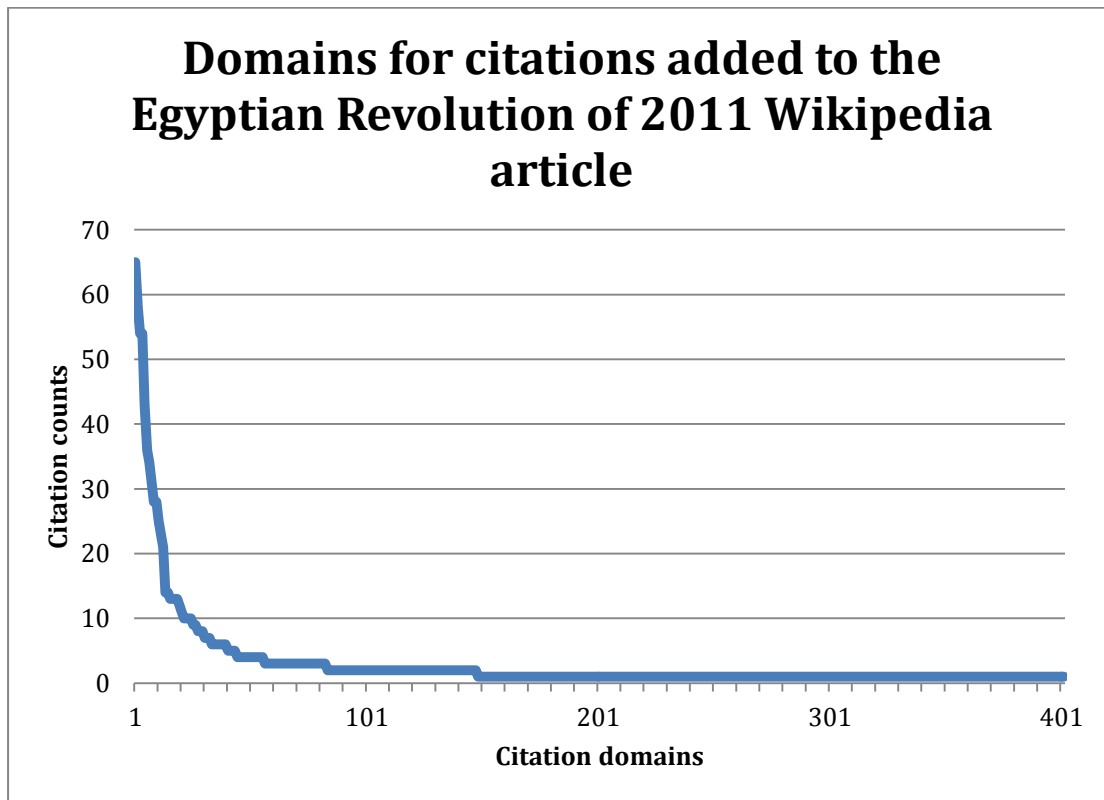
Figure 6.8 Domains of citations added to the Egyptian Revolution of 2011 article from 25 January 2011 to 14 July 2014.

Note: 1,565 citations come from 401 separate domains or sources.

Source: author

Editors had a wide range of sources to select from, although none of these fulfilled the ideal reliable source identified in policy – that of the independent academic analysing events for peer reviewed publications. In the absence of these ideal sources, Wikipedians used what was available: significant quantities of live footage and up-to-the-minute journalistic reports, supported by a network of journalists and citizen journalists tweeting, live blogging and providing regular alerts across a number of news channels.

Journalism

The Egyptian protests became one of the biggest news stories of the year in the United States. The Pew Research Center's Project for Excellence in Journalism News Coverage Index, for example, indicated that coverage of the Egyptian unrest between 31 January and 6 February 'registered as the biggest international story in the past four years—surpassing any coverage of the Iraq war, the Haiti earthquake and the conflict in Afghanistan' (Jurkowitz, 2011). The sheer number of sources available for editors to choose from meant that Egyptian Revolution's traveling companions were therefore much more numerous and diverse in comparison with those found on the surr article. This meant that the Egyptian Revolution was heavily supported in its travel through Wikipedia's infrastructure, and this had a significant impact on how far it was able to travel.

Figure 6.9 represents the 20 most popular citations over time in relation to their numbers after 3.5 years of the article's development. As the article developed, sources and citations were added and removed by editors as a result of a number of factors explained in the section below. Over time, the article became more stable as fewer and fewer edits were made, with fewer citations added and removed.

20 Most popular sources in the Egyptian Revolution of 2011 article
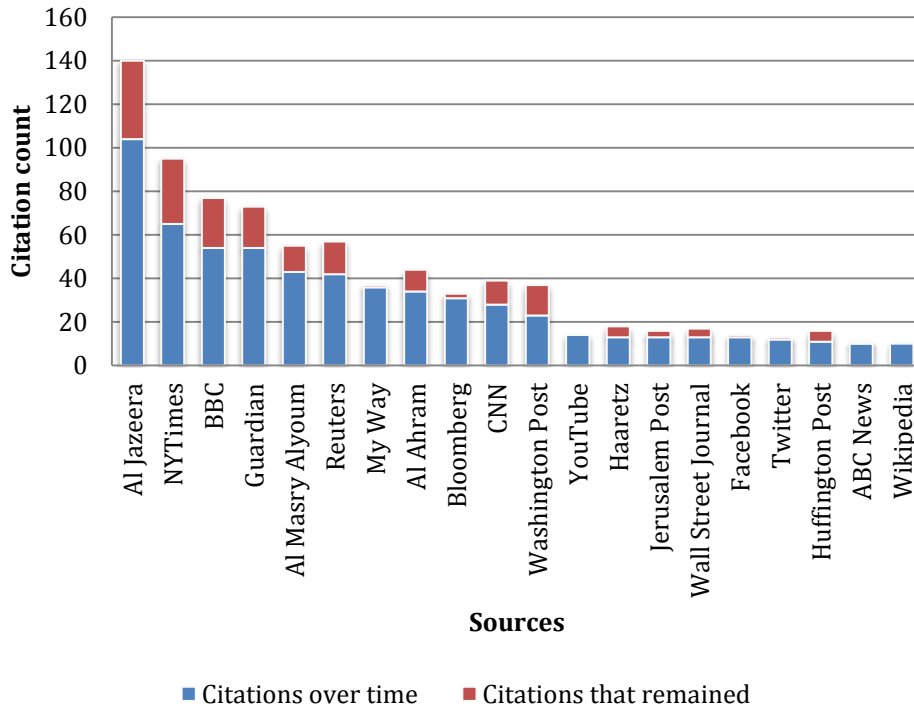
Figure 6.9 20 Most popular sources in the 2011 Egyptian Revolution article

Note: Citations were compared between when the article was first published on 25 January 2011 and when the data was collected on 14 July 2014. A total of 1,565 citations were published over time and 500 citations remained when data was collected on 14 July 2014

Source: author

In the 14 July 2014 version of the article, 500 citations remained. After removing 112 archived sources and 63 broken links, the remaining 325 sources were sources predominantly created by journalists, but there were also a small proportion of citations created by organisations such as Human Rights Watch and Amnesty International, by academics writing for online publications, bloggers or citizen journalists providing individual analysis, by organisations and groups curating directories or archives of resources related to the revolution, documentary films documenting life in Egypt before,

during and after the revolution, as well as two citations from social media (see figure

6.10 below).

Authors of citations for Egyptian Revolution of 2011 Wikipedia article
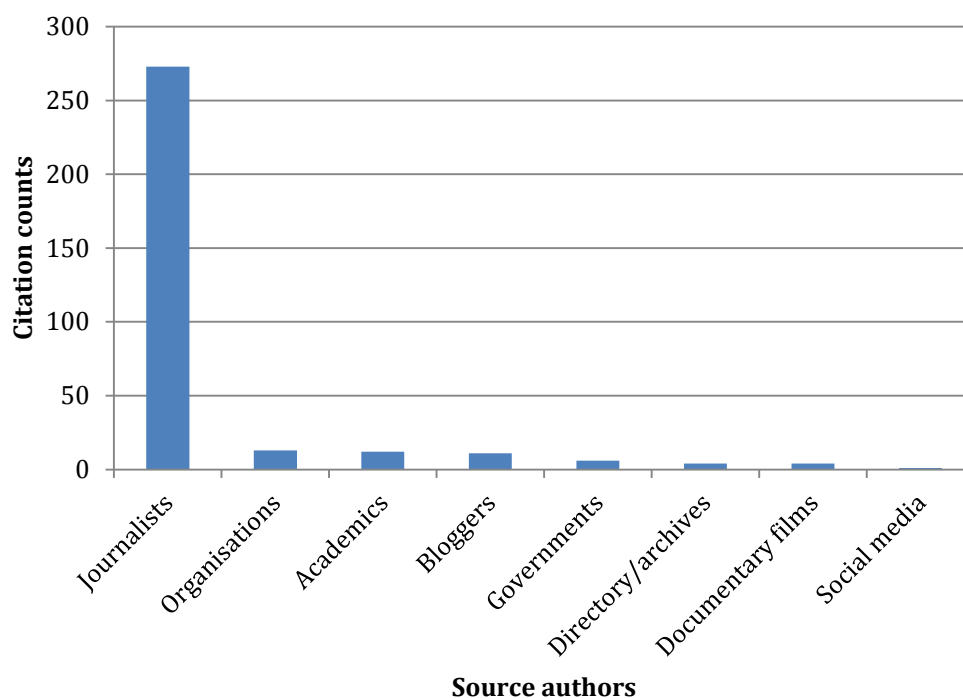
Figure 6.10 Citations for the Egyptian Revolution of 2011 English Wikipedia article according to author type

Note: More information on how classification of sources was conducted is available in chapter three.

Source: author

The most prominent source among journalistic citations was to Al Jazeera English[48] (see

figure 6.11 below).

---

[48] References from now on are to Al Jazeera English

Most popular media sources from the 2011 Egyptian Revolution Wikipedia article
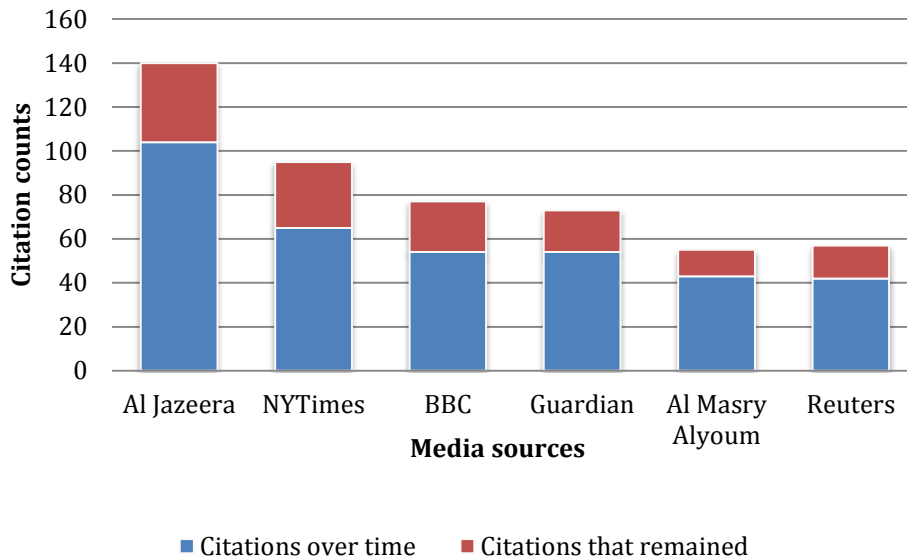
Figure 6.11 Most popular media sources from the 2011 Egyptian Revolution article

Note: Citations were compared between when the article was first published on 25 January 2011 and when the data was collected on 14 July 2014. A total of 1,565 citations were published over time and 500 citations remained when data was collected on 14 July 2014. Citations are predominantly to the English versions of the publications.

Source: author

When editors made statements affirming Al Jazeera as a reliable source in the talk page of the article, other editors did not oppose them, and Al Jazeera stories were regularly cited on the talk page when editors updated one another about what was happening. On the 6th of February, for example, [Lihaas]'s edits claiming that then-President Hosni Mubarak had resigned from his party were reverted by user, [Diaa abdelmoneim]. [Lihaas] opposed the move by stating that 'al jazeera (which has been unanimously declared a fair source above) mention(s) his resignation'. [Diaa abdelmoneim] did not contradict [Lihaas]'s claim that Al Jazeera was a reliable source but retorted: 'This information might have appeared on Al Jazeera briefly but it simply isn't true!!!'

225

(Egyptian Revolution of 2011 talk page, 6-7 February 2011)

Al Jazeera's regular, multichannel, multilingual and openly licensed approach fitted well with Wikipedians' practice of information seeking and article construction. Although there was no widespread agreement on Al Jazeera as a reliable source before the revolution on Wikipedia, a number of editors of the Egyptian Revolution article noted that the media channel gained this status by regular, accessible, textual accounts. According to [Ocaasi], Al Jazeera was 'leading the breaking news (and) would be the first to report' ([Ocaasi] interview), news stories on the website using footage as well as textual analysis, and a live blog providing short textual analysis of events as they happened.

The Egyptian Revolution signalled an important moment in Al Jazeera's global rise to prominence, especially in the United States where it had been banned by cable television companies and declared as anti-American during the Iraq war. In the first days of the protests, Al Jazeera English's website saw an increase in traffic by 2,500%, 60% of which originated from the United States (Stelter, 2011). News reports at the time indicated that the US government was paying attention to Al Jazeera and that US State Department officials including Hilary Clinton had commented that Al Jazeera was providing 'real' news that US stations were failing to deliver (Folkenflik, 2011).

Coverage by Al Jazeera was unique because it was live, round-the-clock and featured up-close images that other news agencies did not have access to. Al Jazeera was also the first international news organisation on the scene since it had offices in Egypt. Americans, in particular, the highest proportion of Wikipedia English editors, were particularly enamoured by their coverage as the interview data further in the chapter and aforementioned media articles indicate. Al Jazeera also agreed to license their

images under a Creative Commons license that would be compatible with their being published on Wikipedia. In numerous ways, then, Al Jazeera suited the format, practices and policies of Wikipedia in the early stages of the article's development.

Academia

Although there were few academics writing in peer reviewed publications during the first few weeks of the events in Egypt, since then there has been a wave of academic research about Egypt's revolution and its role within the Arab Spring. Despite this, only a small proportion of citations within the article can be classified as those originating from the academy.

Only 25 citations relate to academic authors or publishers, the majority of which are to books and peer-reviewed journals (see figure 6.12) and all but one of these journals are openly accessible. The majority of academic citations occur in the further reading section of the article rather than the main body, indicating that academic sources were primary used to guide readers to further reading on the Egyptian Revolution rather than to support statements within the main body of the article.

Academic citations from the Egyptian Revolution of 2011 article
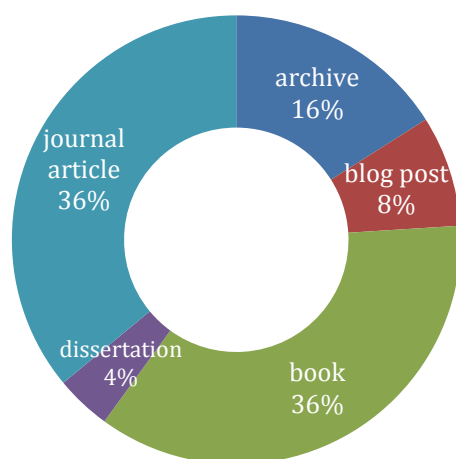
Figure 6.12 Academic citations from the Egyptian Revolution of 2011 Wikipedia article as of 14 July 2014

Note: n=25

Source: author

Almost all of the academic sources cited in the article are clearly aimed at the general public in addition to the scholarly audience. One citation is from the Social Science Research Council, which set up a dedicated website and commissioned online essays from academics such as the author of cited article, Professor of Politics at Georgetown University, Samer Shehata. The essay explores Egyptian perceptions of the United States since 9/11 and was used to cite a phrase in the section on Egypt's emergency law of 1958. The essay is less than 4,000 words, and although it contains footnotes and citations, is formatted according to a main body and conclusion rather than with sections for methodology and results that is common in articles targeted at an academic audience.

In addition to more traditional academic-style articles, there are also two examples of blog posts, one by an Internet technology analyst who has a PhD in computer science but works for a technology firm, another by a Master's student in a student-run online publication. The American University in Cairo also hosts a series of repositories about the Egyptian Revolution and the Arab Spring that were used as citations in the article. One of the sites that was cited is dedicated specifically to the work of faculty and students, the other is an archive for 'photos, videos, visual art, and oral histories contributed by student activists, academics, security officers, and demonstrators in and around Cairo' (American University in Cairo, n.d.). Almost half of the academic-related citations are contained in the 'further reading' section of the article, four from the analysis section relating to online activism, social media and art, one from 'background' and the other from the 'prelude' sections.

The sources that gained prominence in the article, as well as the positioning of those sources, is the result of the particularities of Wikipedia's infrastructure and the practices that editors engage with in editing articles. These factors that determine the fact's terrain make it easier for certain facts to travel through Wikipedia than others since their traveling companions are seen to be more authoritative. In this case, the terrain suited sources that were accessible according to Wikipedians' source practices. Editors still needed to use their knowledge of the Wikipedia environment in order to deploy objects and use appropriate speech acts strategically, but the terrain made the job of some editors easier than others.

6.3   True companions

In addition to the visible traveling companions for the facts of the 2011 Egyptian Revolution, a host of other companions that proved to be influential to the travel of certain facts are revealed through the analysis of editing practice. The true sources of facts in the article were laid bare when editors revealed their practices, both in the talk page and through interviews with editors.

Some of the sources that were added to the article were more frequently removed by editors over time. Al Jazeera, social media sources and Google were less visible after the article had stabilized than while the article was in development. While Al Jazeera features prominently in the article's citations, it was also frequently removed and replaced with sources such as New York Times and the BBC. Social media sources were almost entirely removed over time. Sources that were unacknowledged as the source of facts in the article but that were used to discover and verify facts contained within it included Google and Wikipedians who were in Egypt at the time.

6.3.1   Google

Google was prominently discussed in the talk page of the article as a source of news and trends, but very few citations were to Google and those that appeared were quickly removed. Google News was used by editors as a channel for receiving new facts about the Egyptian Revolution. [Ocaasi], for example, noted that he was regularly watching Al Jazeera English's live television coverage and used Google to find stories in online newspapers when new events were flagged up by the live coverage.

> There were different streams of immediacy... well, I was one of them just glued to Al Jazeera's live feed from Tahrir Square, and so I was watching that and then when they would say something new, then I would go and look for whether or not Associated Press or New York Times had run with anything, so I wouldn't use... I would kind of use the breaking news television as inspiration to go find sources and then at the same time I was just scouring Google News for the

information coming in about the Egyptian Revolution. ([Ocaasi], Skype interview, 3 April 2014)

[SilverSerenC] who took on the role of keeping editors apprised of new sources on the talk page of the article recalled how he used Google News to keep updated with the latest sources.

> I generally use Google News and do specific keyword searches, or search for news in the past 24 hours. Every hour that I was up, I would check the most recent news sources for things that should probably be put in the article. ([SilverSerenC], Skype interview, 22 August 2012)

In accounts of editing and sourcing practice, editors noted that their process began with watching live video and encountering social media accounts of what was happening on the ground and that they used Google to search for alternative (or additional) sources of information in the article. Video sources are difficult to cite because they are not always archived, and when they are archived, editors must be able to discover where the archives of the live footage are housed. Waiting for video archives to be created and then discovering the right footage is a much more lengthy process than using a Google News search in order to find textual news stories to read and cite.

Editors also regularly used Google Trends reports to argue for different titles of the article. Different sources used different names at different times and each day of the protest had its own title. The 28th of January 2011, for example, was called the 'Friday of Anger' by some and as the 'Day of Rage by others. The article was created with the title '2011 Egyptian protests', but after a few days, some editors called for changes to the primary name of the article, primarily calling for a change to either 'uprising' or 'revolution'.

Changing the primary title of the article had important political consequences, and naming became one of the most discussed issues on the talk page. Article titles

constitute the primary facts of the article and carry enormous weight in shaping the secondary facts that are contained within it. Whether events in Egypt constituted a 'Revolution' or mere 'protests', for example, had implications for the way in which events were considered by the Egyptian state as well as international onlookers.

Editors argued that the term 'protests' could mean 'any isolated events', but the term 'uprising' denoted a much more extensive national phenomenon and the term 'revolution' implied not only widespread legitimacy but also success. In a discussion about the use of the term 'protests' or 'uprising', [Ocaasi] noted the different numbers of citations to the terms in terms of their rough numbers:

> Just a rough guide, Google News search limited to the past 7 days for **Egypt protests**: *27,693*; for **Egypt uprising**: *11,305* (note: similar breakdown of about 2.5:1 for egyptian protests, egyptian uprising). See also Google Trends (egypt protests, egypt uprising) [1]. And check out the Google Trends regional breakout at the bottom, where the 'uprising' term has indeed caught on more in the US than the UK or Canada:[2] ([Ocaasi], Egyptian Revolution of 2011 talk page, 1 February 2011)

Other editors, including [Lothar von Richthofen] and [Userpd] provided links to Google news searches for the alternative terms as evidence for their votes in opposition or support but [Cs32en] countered that editors should rather look to what reliable sources had been using in their reports on events since Google searches aren't the most accurate method of determining what the events in Egypt were being called.

> (A) Google search can't really answer the question of whether *the movement as a whole* is being seen as an uprising, a revolution, a revolt, or as (a number of) protests. We have to look carefully at the context in which reliable sources use words such as "protests" or "uprising"... (Cs32en, Egyptian Revolution of 2011 talk page, 2 February 2011)

In another response, [Userpd] demonstrated that not all editors had the same understanding of what sources were considered reliable outside of those regularly cited on English Wikipedia such as the BBC, the New York Times and the Guardian. In the

232

comment below [Userpd] opposed the nominator's suggestion that the article title be changed from 'protests' to 'uprising'. [Userpd] wrote that s/he believes the nominator to be 'known for his pro-right Israeli edits' and implied that it is in Israel's best interests to discredit the protesters. Russian sources, on the other hand, had been calling the events 'protests' from 'the beginning'.

> ...in other sources in different languages the word "Protest" (was used from the beginning), for example russian ones: gazeta.ru, lenta.ru, 1 state channel, Russia. So I think we can manage without this cliche / label from mainly pro-right Israeli side. ([Userpd], Egyptian Revolution of 2011 talk page, 6 February 2011)

[SilverSerenC] responded with his own interpretation of Russian sources, asserting that it is actually in the *Russian* media's best interests to call events in Egypt a protest rather than an uprising because they fear similar actions by 'former bloc countries'.

> As for Russia, (I)'m not sure what you're expecting to find there. Of course they don't use the term Uprising. That is a term that Russian media does not want to go throwing around. Russia might not be being as extravagantly censorist as China is, but they are still being careful of what they say. Not that Russia has to worry about that as much from the main population, they love Russia, but more from the outlying territories, the former bloc countries. Believe me, I know how Russia works. Go read my article on the <u>War of Laws</u>. ([SilverSerenC], Egyptian Revolution of 2011, 6 February 2011)

In the context of such complicated weighing of the interests of each source implied by these two contrasting statements, calculating the numerical weight of each name according to Google starts to seem a much simpler mechanism. Google, consequently, became an obligatory passage point in the construction of the article, both because editors used Google News to search for updates to what had happened and to corroborate or refute evidence found elsewhere, and also because Google Trends was used as a source for deciding the politically important article names.

Google does not represent all information sources about a topic; its algorithm is determined by rules that are driven by the goals set by Google's directors and investors

233

(Gillespie, 2014). Google News only indexes websites that allow indexing by Google and Google Trends only illustrates terms that show up in the corpus that it has indexed through Google Books. Furthermore, other facts of the revolution, including the discussions and negotiations that were undertaken behind closed doors in the lead up to Mubarak's resignation and its aftermath, were simply unavailable anywhere online.

Coverage by sources on the Internet was only a small proportion of what events consisted of, even though many editors believed that what was available online represented the universe of all there was to know about the event. Editors regularly cited numbers of Google results for different terms, implying that they if they read everything available via Google, for example, they would know what the events were being called as a whole; if they calculated how many times the different terms were being used, they could determine what the true name was. Since the representation on Wikipedia does not reflect the entirety of events that took place in Egypt in 2011, however, the consequence of editors using Google in order to make editorial decisions was that Google was acting as a filter for sources of information on Wikipedia that remain hidden to the readers of the article.

6.3.2   Social media

Social media sources that were cited or mentioned in talk pages include the Facebook groups used to organize protests, the Twitter accounts of Egyptians and the blogs of those active in the protests. Interviews with editors and conversations on the talk page of the Egyptian Revolution article indicate that social media streams became particularly important sources of information for editors, even if their citations did not reflect this.

Some editors noted that social media sources were used by them in order to discover

new information related to events and that such awareness catalysed deeper research from more reliable sources. There were, however, cases in which lines between social media and traditional media were blurred by editors citing facts.

> I think social media... it inspired research. You know, it made you look for confirmation for things, but in general, just because something was Tweeted, especially if it was just Tweeted by an Egyptian protestor, that wouldn't be sufficient; we would never cite that. If a journalist who was on the ground in Egypt Tweeted something then... again, it still exists in this grey area; we would tend to want to wait for a full article to come out to confirm it, but there were definitely... there was a bit of blurring of the lines. ([Ocaasi], Skype interview, 3 April 2014)

Social media sources were removed not only because they are not considered ideal in terms of policy but also because of problems relating to their archival. [Aude] recalled how she was following who she believed were reliable journalists on Twitter but that she preferred to replace Twitter citations with 'something more permanent'.

> And then you'd have some other reporters like from Al Jazeera, they were Tweeting and so if you (thought that) 'X' person was reliable... their tweets could be reliable, but in a pinch you could use something like that, as a source, but when you have the time to go back and fix up the article, it's good to find something that's more permanent. ([Aude], Skype interview, 4 May 2012)

Twitter no longer archives all Tweets of its users which means that there is often no permanent URL on Twitter for editors to cite. Other services (such as archive.org) sometimes archive Tweets that they believe capture important historical moments and editors can use free services to archive web pages in order to attach a permanent URL to the sources that they cite. 12% of the citations used in the article were to URLs that were no longer available when the data was downloaded which indicates that few knew about archiving of URLs or think that it is too time consuming to archive a source before citing it on Wikipedia.

An additional problem faced by editors was that YouTube videos originally cited in the article appeared to be copyright violations and thus had to be removed. The only two

remaining social media sources as of 14 July 2014 were a Twitter post from Mohamed
ElBaradei from 22 November 2012 (figure 6.13) and a citation to Facebook after a
phrase about the numbers of members (325,000) of the 6 April Youth Movement
Facebook group.
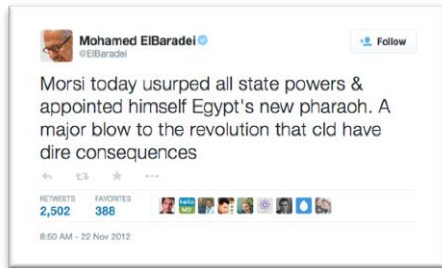
Figure 6.13 Mohamed ElBaradei's tweet linked as a citation to the Egyptian Revolution of 2011 article as of 14 July 2014

Source: https://twitter.com/ElBaradei/status/271656968341581824
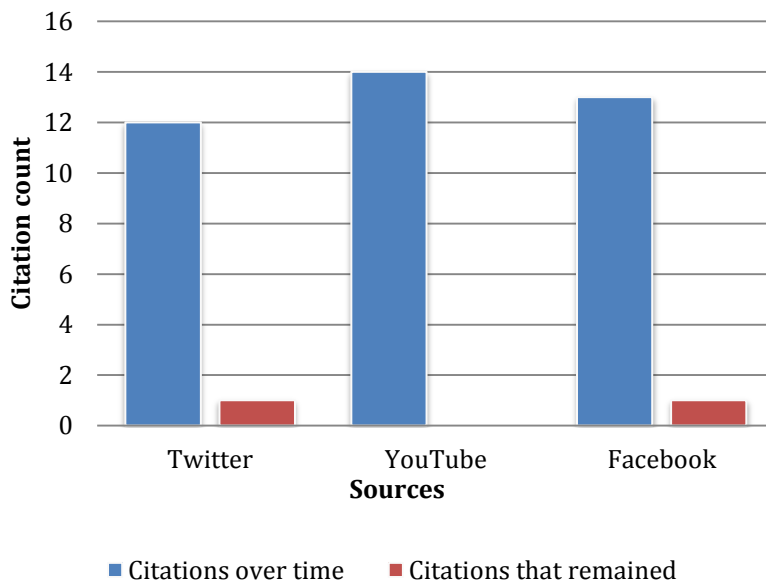
Figure 6.14 Citation counts of social media sources removed over time from Egyptian Revolution of 2011 article

Note: Citations were compared between when the article was first published on 25 January 2011 and when the data was collected on 14 July 2014. A total of 1,565 citations were published over time and 500 citations remained when data was collected on 14 July 2014.

Source: https://en.wikipedia.org/w/api.php

Figure 6.14 depicts the proportions of sources from Twitter, YouTube and Facebook that were added over time, and the two sources that remained when data was collected on 14 July 2014. Although social media sources were highlighted as one of the primary ways in which editors discovered news and information about the events of 2011, bias against social media sources in terms of both policy and technical difficulties resulted in a significantly steeper rate of removal for social media sources than traditional media sources as can be seen in figure 6.15.

Rate at which traditional media vs. social media citations were removed from the Egyptian Revolution of 2011 article
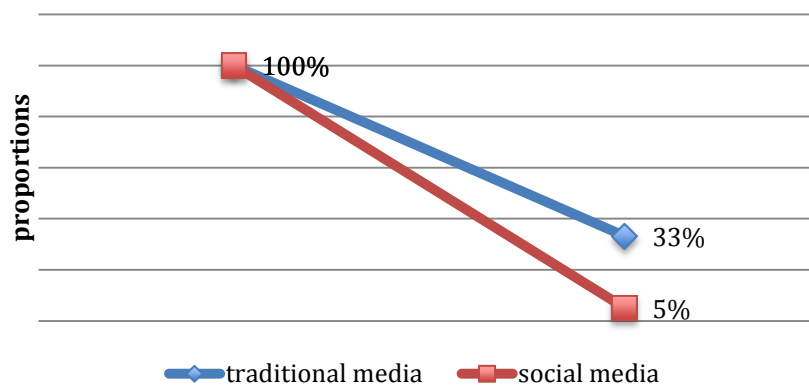
Figure 6.15 Rate at which traditional media vs. social media citations were removed from the Egyptian Revolution of 2011 Wikipedia article

Note: Citations were compared between when the article was first published on 25 January 2011 and when the data was collected on 14 July 2014. A total of 1,565 citations were published over time and 500 citations remained when data was collected on 14 July 2014.

Source: https://en.wikipedia.org/w/api.php

Social media sources were deemed useful to editors in some cases, such as when Facebook was used to determine which names were associated with which days during protests. In one case, editors were confused between the 'Day of Anger', the 'Friday of Anger' and the 'Day of Rage'. [Lihaas] and [The Egyptian Liberal] confirmed that these were all separate events, with [The Egyptian Liberal] pointing to the Facebook page where the 'Friday of Anger' name originated from.

> **[112.119.91.68]:** It seems that 'Day of Anger' is much more common than 'Friday of Anger'. That's what Al Jazeera English is using, and dominates in Google hits. Any comment?
>
> But 'Day of Rage' (which the BBC is using) gets just about the same number of G-hits as 'Day of Anger'.
>
> **[Michaelzeng7]:** @[The Egyptian Liberal] Could you link to that facebook maybe?
>
> **[Lihaas]:** day of anger is a *different* day from yesterdays events. (id guess something that failed)
>
> **[The Egyptian Liberal]:** Day of Anger is the 25th ... Friday of Anger is the 28th ... and here is the link to the event page on facebook [6]
>
> (Egyptian Revolution of 2011 talk page, 29 January 2011)

Despite the use of the Facebook source as a way of verifying the titles of protests, the citation was removed shortly afterwards[49]. The sections of the article dealing with the role of Facebook and online activism in the protests have only a single link to the 6 April Youth Movement Facebook group mentioned earlier. The Facebook group, 'We are all Khaled Saeed[50]' that catalysed support for the first protests on 25 January 2011, for example, is not cited. Instead, there are multiple citations to analysis of the role of social media platforms in the revolution from sites such as The New York Times online, the

---

[49] The URL is also no longer available on Facebook.

[50] https://www.facebook.com/elshaheeed.co.uk.

Guardian online and the Atlantic. Wikipedians have removed primary sources, or sources that enable the demonstration of particular facts, in favour of chiefly indicating secondary sources.

Social media sources acted as the primary channels through which the subjects of the article were able to have their voices heard. Facebook groups that were used to help organize protests, the Twitter accounts of Egyptians and the blogs of those active in the protests and the footage taken by people during protests were the primary media through which the subjects of the Egyptian Revolution of 2011 article could speak. The iterative removal of these artifacts constituted an iterative exclusion of the voices of the subject from the article.

6.3.3 Wikipedians

Another set of the Egyptian Revolutions traveling companions appeared in the evidence provided by Wikipedians themselves. Although editors are urged from contributing their own knowledge to the article through the policy of 'no original research', there are numerous occasions in which the line between editorial contributions that are allowed (such as summarizing sources) and those that are not allowed (such as adding facts from their own knowledge rather than cited to independent sources) becomes blurred.

One such example occurred when editors were discussing the Google employee, Wael Ghonim who started the Facebook group, 'We are all Khaled Saeed' that called for protests on the 25th of Janury. Editors discussed whether the Wall Street Journal's claim that Ghonim was to be released was true and whether he would still be the administrator of the Facebook Group.

> **[The Egyptian Liberal]:** I just spoke to Wael and no, he is still the admin of the page
>
> (Egyptian Revolution of 2011 talk page, 7-8 February 2011)

[The Egyptian Liberal] was in Egypt at the time and had personal connections to protesters and activists in Egypt, and was thus able to verify that Wael Ghonim was still the administrator of the Facebook group. [The Egyptian Liberal] is, however, unacknowledged as a source in the article.

In another conversation, the contribution of Wikipedia editors who were also participants of the protests was subject to debate. Three days after the first protests, one editor requested that unverified accounts by protesters be allowed in the article as a means for protesters to be able to communicate. Other editors responded by reaffirming the independence of the encyclopaedia and by providing advice to those involved about how they might be able to influence what is written in the article.

> **[Athinker]:** ALLOW UNVERIFIED COMMENTS IN THE ARTICLE AND LABEL IT SO: SOME PROTESTERS CANNOT COMMUNICATE OTHERWISE
>
> **[SilverSerenC]:** They have this talk page. The article is for (encyclopaedic), verifiable information, not unverified comments from protesters.

Editors [Edison] and [Wnt] then provided advice on how editors could become influential in the article by becoming sources in news stories that Wikipedia could then quote and cite.

> **[Edison]:** Wikipedia is absolutely not here to provide a forum or soapbox for peoples' views on the situation. Protesters can talk to news service reporters on site in Egypt, such as Al Jazeera and we can reference the news service articles if they choose to report them. Land line service is also said to be in service from Egypt to the outside, allowing spokesmen for the protesters to contact news services.
>
> **[Wnt]:** Wikipedia's purpose is to cover the event in detail. Technically, the encyclopedia is neutral - though if one side in a conflict pursues censorship, we

240

will inevitably tend to undermine that effort. Facts from involved participants are always welcome on this talk page. After all, it's a lot easier to find a source for something if you know what that something is. Unfortunately, we can't add facts straight from a random Wikipedia account unless they are verifiable to some kind of reliable source.

[Wnt] then outlined three ways that protesters could become involved by highlighting the avenues available to individuals to influence Wikipedia content. These included contributing to Wikimedia Commons (which s/he believed allows editors to create their own content), creating an external source which can be referred to, or by adding their content to Wikinews (another Wikimedia project that hosts original news reports written by individuals).

> **[Wnt]:** Some possible "loopholes" do exist for people involved in the protests to get facts into the article without violating Wikipedia policy. These include:
>
> **Submit a photo** to Wikimedia Commons. We need public domain photos of these historic events. When you submit a photo, you can include a detailed annotation about what was happening and when. You can provide a shorter description in the caption for the photo when it is placed into this article. This may be usable even for less dramatic events. For example, if you find that Wikipedia is blocked in Egypt, but that you can get around it with an open proxy, you can't just write here that you can do this because it would be original research - but you *can* post photos showing your computer windows, one blocked, one getting through, and it might make it into the article (I predict debate, though)
>
> **Create or join a partisan website**. Wikipedia can report on partisan primary sources that are relatively notable, so if you put something up on the Web that looks like a credible statement from a group of people, it can be summarized here as one side's opinion.
>
> **Use Wikinews.** That site welcomes journalistic reports. I am not very familiar with it, but you may find it a very useful way to get the story out.
>
> (Egyptian Revolution of 2011 article talk page, 28-29 January 2011)

The line between an editor engaging in active construction and passive recording is difficult for editors to navigate. In the surr example, we saw how editors were forbidden from applying their own knowledge to the development of the article, but in this case, editors' knowledge was applied in ways that were tacitly accepted by other editors. All

knowledge work requires that knowledge workers apply their own knowledge to choices about how facts should be summarised and supported, which elements of the narrative are emphasized over others, and which views are given dominant status and which are designed as fringe views. By denying this feature of knowledge work, editors must effectively hide their contributions or expel those who are explicit about their interests.

The difference between surr and the Egyptian Revolution cases was that in the case of the Egyptian Revolution, editors were successfully able to apply their knowledge of Wikipedia's norms and policies in order to frame their activities as in keeping with policies around original research. Those who were new to Wikipedia, represented by those who asked to allow unverified comments in the article, for example, were relegated to using alternative media in order to have their voices heard.

6.4   How well did the fact travel?

The facts of the Egyptian Revolution of 2011 Wikipedia English article travelled beyond its initial boundaries to find new users and uses. The article and its attendant facts exist in 58 language versions of the encyclopaedia and almost 2,000 articles and pages within Wikipedia link to the article, with at least eight new articles started to extend themes established in the article. The article is also well represented in Wikidata (see figure 6.16), where it provides links to all the language versions of Wikipedia in which the article occurs, links to the images in Wikimedia Commons created under the heading 'Egyptian Revolution of 2011' as well as links to other Wikimedia Projects in which the

Egyptian Revolution facts occur such as WikiQuote which contains famous quotations from the events in five languages.[51]

Statements about the location, categories and dates related to the event in Wikidata serve to provide factual information that enables the Egyptian Revolution facts contained within Wikipedia to find further users and uses when other online services use this data to populate their own websites. Wikidata's centralization of facts relating to the 2011 Egyptian Revolution throughout all Wikimedia projects also helps to enable greater integrity of the facts as it becomes simpler to make facts across the sites more consistent.
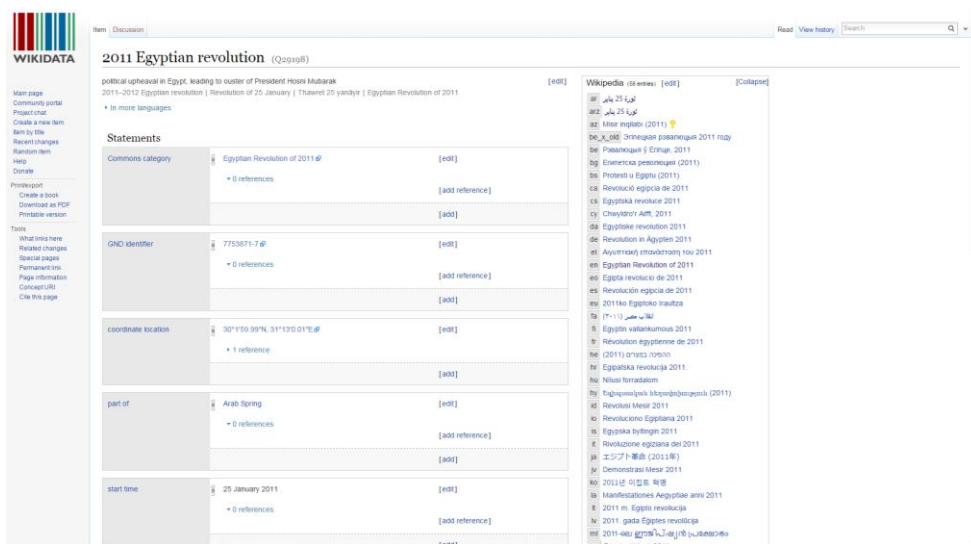
Figure 6.16 The 2011 Egyptian Revolution as represented in Wikidata as Q29198

Source: https://www.wikidata.org/wiki/Q1218 as of 1 August 2015

The success of the Egyptian Revolution facts on English Wikipedia in traveling well was partly due to the facts' traveling companions. Traveling companions consisted of the editors who shaped and supported the facts, the sources (both visible and invisible) that

---

[51] See https://en.wikiquote.org/wiki/2011_Egyptian_revolution.

editors used to construct or verify facts as well as the facts' labels. Labelling was achieved in numerous ways, using warning tags, categories, vandalism detection and data extraction by both automated and non-automated means. Bots enabled the Egyptian Revolution article to enter Wikidata, and Wikidata editors added statements to the Egyptian Revolution item.

Traveling companions for the Egyptian Revolution facts were numerous. The fact was surprising because it was unexpected; it also represented a narrative (ordinary people rising up to unseat their dictator in what appeared to be a rise of democratic solidarity) that excited people and their governments in Western nations. It was not enough, however, for the facts' allies to be numerous. Allies also had to be deemed legitimate and reliable by the editors who were choosing them and by the readers in whose interests they were being chosen.

Sources of facts contained within the article were either visible to other editors but invisible to readers; or visible to editors and to readers when they took the form of citations, especially those that remained stable over time. Sources that were influential were those that were easily accessible, citeable and were not heavily challenged. Some sources' visibility decayed over time so that they were not displayed as prominently as they were when the article was originally constructed. Other sources were visible only to editors or groups of editors, rather than readers, because they were not attributed. As Morgan notes, 'what appears to be freely travelling knowledge is in fact dependent on tacit or expert community knowledge to make it transfer effectively' (Morgan, 2010b, p. 26). Facts often require invisible allies that take the form of the tacit knowledge required to know how to present, shape and move facts along certain paths.

Facts offload some of their traveling companions as they become more accepted and thus no longer need allies and can stand on their own. When the article was renamed to 'Revolution' replacing 'protests', for example, it was accompanied by numerous citations, but these were removed as the name became more well-known and acceptable.

Elements of the terrain that proved important for the travel of the Egyptian Revolution facts included policy, practice and materiality. The sources that were available to support the Egyptian Revolution facts' travel in the first stages of the events were primarily media and social media sources, sources that did not fit the ideal of academic sources in books and journals that Wikipedia policy prescribes. The practices of editors also determined which facts and their traveling companions were chosen. Sources that suited Wikipedians' 24/7 editing practices and the requirement for citations to support each fact or statement meant that Al Jazeera was the most popular source. Finally, the materiality of sources was important because they needed to be archived and accessible, both in style and in terms of price (that is free, rather than closed access for a fee).

In each case, the terrain needed to be shaped by knowledgeable editors in order to prove advantageous to the facts' travel. Experienced editors had to choose the most reliable sources according to their perceptions of what readers and other editors considered reliable. Editors had to defend their choice of source as reliable (or their designation of a source as unreliable) by using appropriate rhetoric that included the deployment of policy objects and disciplinary processes. Editors also had to be aware of archiving procedures in order to sustain citations over time.

Facts travelled within a terrain that was influenced by what was acceptable according to Wikipedia policy, the practices of editors as they discovered, verified and summarised sources, and the materiality of available sources which suited Wikipedia editing practice. The Egyptian Revolution facts also travelled well because of the character of the fact itself. The appearance of thousands of ordinary people in Tahrir Square was especially interesting to people in the West, as it symbolised a democratic movement against authoritarianism. Events in Egypt registered as the biggest international story in the United States media in the previous four years as indicated earlier and the media attention helped to catalyse the rapid travel of the fact beyond its first establishment as an article on English Wikipedia.

6.5   Conclusion

The implications for how facts travel within Wikipedia's socio-technical system are that certain identities are accorded with greater authority than others, and therefore have greater power to determine the representation of facts on Wikipedia than others. Two sources of authority were enacted in the progress of the Egyptian Revolution of 2011 article. The first in the authority of certain information sources above others, and the second in the authority of certain editors above others.

Revealing the true companions of the facts that make up the 2011 Egyptian Revolution English article indicates that there are key factors that affect which traveling companions are revealed and which remain hidden in Wikipedia. Firstly, the materiality of sources is important, particularly in relation to whether sources are accessible, archived and/or archivable. Accessibility consists both of the readability of scholarly texts by non-academic audiences and the availability of academic and journalistic texts to the public since Wikipedia policy favours the use of texts that all readers can verify.

The extent to which certain sources such as Twitter are archived or archivable is important because, even if sources can be archived using free services, archiving requires the knowledge of which services to use and the time in order to archive sources before citation.

Authoritative editors are also important allies in the travel of facts, although they remain largely hidden from the article. Since norms around acceptability of sources and tacit knowledge about the management and display of sources are so important to having certain facts accepted, an editor's tacit knowledge and ability to actively demonstrate their knowledge becomes a significant factor in having facts accepted. Editors who made the choices about which sources to remove over time were instrumental because they shaped the article, deciding which traveling companions were considered reliable enough and which should be replaced. Experienced editors have the tacit knowledge necessary to either enable the fact to travel further or stop it in its tracks.

There are unexpected consequences of this particular assemblage of policies, editors, practices, norms and procedures at the edges of Wikipedia. Power becomes lodged in discourses that posit knowledge subjects (in this case, the people of Egypt) to be illegitimate sources of knowledge and representations in which the subject of facts were designated as their traveling companions were iteratively removed over time. Finally, power becomes lodged in identities, in particular the identity of the Wikipedian who by virtue of her/his edit tacit knowledge and experience of Wikipedia's policies, norms, practices and procedures is able to determine which traveling companions are determined as legitimate and what the role of the knowledge subject is in the representation of facts.

In the next chapter, the core findings of this thesis are brought together in an analysis of how facts travel in Wikipedia, why some facts travel further than others and what this means for how Wikipedia is reconfiguring the roles and authority of knowledge workers.

Chapter 7: Wikipedia's search for encyclopaedic identity

> When the world one knows is in disarray, redefining identities is a way of putting things back into familiar place. (Jasanoff, 2004, p. 39)

In chapter one of this thesis, the two research questions under analysis were identified and explained.

> RQ1: Can facts be shown to travel within Wikipedia's socio-technical system? If so, do some facts travel further and with greater integrity than others, and why is this so?

> RQ2: Is there evidence that the roles of knowledge authorities and experts are being reconfigured within Wikipedia's socio-technical system, and if so, how has this reconfiguration taken place?

In the past three chapters, these questions have been analysed according to three distinct followings. Chapter four outlined the terrain in which facts travel, indicating the principle of 'verifiability, not truth'. Verifiability was found to be a central priority and value system that guided the early development of Wikipedia, signalling a distinct departure from the epistemological foundations of past encyclopaedias. Chapter five and six followed the travel of two sets of facts through Wikipedia's socio-technical system, uncovering how the facts' traveling companions and character influenced how well they were able to travel. Facts have, indeed, been shown to travel within Wikipedia's socio-technical system, with certain facts traveling further and with greater integrity than others as demonstrated in the different trajectories of the surr and Egyptian Revolution articles (RQ1). Implications for the ways in which power becomes lodged in identities, representations and discourses as a result of the trajectory of facts' travel was introduced in these chapters in response to RQ2 but extended here.

This chapter covers the three core themes of the project: how facts were shown to travel within Wikipedia's socio-technical system and the factors supporting facts' travelling well within that system. The chapter moves on to a discussion of how knowledge authorities and experts are being reconfigured within Wikipedia. Finally, there is an analysis of the implications for the travel of facts more generally as power become lodged in identities, discourses and representations within Wikipedia and the platforms that extract facts originating in Wikipedia. There are three core findings from this thesis that are articulated below.

The first is that every edit of Wikipedia constitutes an articulation of the boundary between Wikipedia and the world in the form of Wikipedia's identity, the role of Wikipedia's authors and the identity of reliable sources. In defining Wikipedia as an encyclopedia, proponents must work to shape content according to normative definitions of what is encyclopedic and to contrast Wikipedia to journalism and academic or scientific projects that share similarities with it. In defining Wikipedians as independent, unconcerned individuals, Wikipedians must define themselves as different from the authors of social media platforms who express their opinions about the world rather than merely summarise existing (and authoritative) opinions. In defining reliable sources as published by people who are not Wikipedians, Wikipedians must contrast reliable sources written by independent, unconcerned individuals with the emotional, advocatory, personal opinions of what they deem to be unreliable sources.

The second is that the articulation of these boundaries takes the form of speech acts that are both symbolic and discursive. Instead of rhetorical strategies that Geertz and Gieryn argue are used by ideologists in order to establish their perspective about boundaries, Wikipedia editors enact their chosen boundaries directly on the article by either

stabilizing or destabilizing it. Such strategies indicate that there is a new vocabulary with which knowledge claims are being made that is highly specialized and unique to Wikipedia's socio-technical environment, and that power is being performed on the edges of the encyclopedia rather than through the technical permissions that different classes of editors are provided with. If power is about the ability to affect the representation of a particular subject, then power is about the ways in which socio-technical relations are performed at the article level rather than through policy or other central decision-making bodies.

Finally, I argue that Wikipedia has become authoritative by a process of reaffirming the authority of traditional experts and in doing so Wikipedia has signaled the rise of new centers of expertise. Wikipedia's content is being used by multiple external players and has become an authoritative source for factual statements about the world. Wikipedia has obtained this authority not by abandoning traditional centers of authority, but by re-affirming those centers of power by preferring academic sources, albeit sources aimed at a more public audience, and by dominant English-language media corporations such as the New York Times and the BBC. In doing so, Wikipedians have gained their own authority so that they are now seen as a quality platform in which to find summarized versions of authoritative sources. Since the process of selecting and placing sources is a highly contextualized one, dependent on specialized skills, there is an increasing distance between effective and ineffective Wikipedia editors, indicating a growing professionalization of Wikipedia editing and a thus a growing rift between different classes of editors.

7.1 Wikipedia's traveling facts

The past two chapters demonstrated how facts can be seen to travel within Wikipedia's socio-technical system. I defined the travel of facts as the extraction and/or linking of facts within different Wikimedia sites and objects, and facts as 'autonomous, short, specific and reliable piece(s) of knowledge' (Morgan, 2010, p. 8). Within Wikipedia, facts are represented by article titles, statements that constitute article content, redirects, categories, language links, images and captions. Facts are also contained within other platforms such as Wikimedia Commons that hosts images, maps, videos, sounds for use in Wikimedia projects, WikiSource that hosts quotations from famous people, books, speeches, films and other 'intellectually interesting materials' and meta-wiki which is used as a central hub for coordinating and organising tasks, projects and events.

Each of these networks consists of different assemblages of people, objects, policies, norms, practices, processes and procedures that become enrolled in the representation of facts. In this thesis, I have focused on following facts within English Wikipedia, but have highlighted how facts travel to and from other networks, including the origin of surr on Hindi Wikipedia and its travel to English and Korean Wikipedias, as well as the origin of the Egyptian Revolution of 2011 article on English Wikipedia and its travel to Arabic and Egyptian Arabic Wikipedias. I have also indicated how, when facts travel to Wikidata, they are summarised by a few structured (coded) statements that enable facts to travel more easily and with greater integrity within Wikimedia's numerous sites, and to sites beyond Wikimedia that make use of Wikipedia and other Wikimedia site data such as search engines and other applications.

Whether facts travel well or not depends on their traveling companions, their terrain and their character. I employed Morgan's (2010b) definition of 'traveling well' to

encompass the concepts of integrity and fruitfulness, where integrity was determined by the fact traveling more or less intact from one place to another, and fruitfulness determined by a fact traveling beyond its spatial or disciplinary boundaries to find new users and uses. Travelling companions in Wikipedia's case consists of editors, sources and labels which include tags, categories and links.[52]

The two facts that were followed demonstrated very different trajectories. The facts relating to the surr case travelled a relatively short distance within Wikipedia's socio-technical system since surr is only available in three languages and contains few links to and from the article. Surr contains no statements in Wikidata and links to no other Wikimedia projects. The Egyptian Revolution of 2011, on the other hand, travelled far and wide within Wikipedia and the Wikimedia network more generally. The article exists in 58 language versions of Wikipedia, almost 2,000 articles and pages within Wikipedia link to the article, with at least eight new articles started to extend themes established in the article. The Egyptian Revolution of 2011 is also well represented in Wikidata, WikiQuote and Wikimedia Commons. While surr did not travel with great integrity because of the conflicting statements about who plays surr in the articles and sources, the Egyptian Revolution of 2011 travelled with greater integrity because of the centralization afforded by Wikidata.[53]

---

[52] Categories and links can also be considered to be facts, depending on whether one is analysing the entire article and the labelling of it in the categories and links attached to the article as a whole, or whether analysis is applied to each of the elements within the article in which case the categorisation of a particular phenomenon is considered as a fact relating to that phenomenon.

[53] Further research is required in order to understand the extent to which facts represented on different language versions of the article were similar or different in order to fully understand how intact facts travelled between different language versions.

A number of factors were demonstrated to influence how well facts travel within Wikipedia. These include the extent to which the facts' traveling companions were considered authoritative or not, the character of facts, particularly in relation whether they were interesting to the types of editors who dominate English Wikipedia in the United States and Western Europe. Finally, the facts' terrain was influential in how far the fact travelled. A fact's terrain can be encapsulated by three core themes: the processes in which objects (including policy objects and technological tools) were deployed, the market and materiality of sources available, and the ideologies, discourses and identity of Wikipedia editors.

7.1.1 Traveling companions

In the surr case, surr's traveling companions were considered illegitimate by editors who were opposed to the use of oral citations as designed by Prabhala and his allies. Opponents were able to successfully label oral citations sources as illegitimate through the responses and non-responses of the project's proponents. Proponents effectively acceded to the authority of editors such as [Fifelfoo] by virtue both of their lack of involvement in discussions on the RS/N that proved to be the most important venue for deciding the reliability of all the sources created as part of the project, and by virtue of [Utcursch]'s addition of an online newspaper source about oral citations in response to repeated opposition to oral citations. The designation of the oral citations sources as primary sources by [Sir Nicholas] and non-compliant in terms of copyright standards by [Fifelfoo] and [MER-C] meant that the entire article became labelled as unstable since it initially relied only on the oral citations as sources. Editors indicated the instability of the article by applying warning tags and deleting the content of the article, thereby alerting readers to the unreliability of the facts housed within it.

Editors participating in the RS/N and the Wikimedia-l discussion also labelled the *authors* of the oral citations as illegitimate. In the RS/N discussion, interviewers and interviewees represented in the oral citations were labelled as 'amateur' and 'unreliable' as opposed to 'professional' or 'academic'. The authority of both the interviewer and interviewee to speak on behalf of the subject of the article was brought into question, but proponents did not respond to such labelling on the RS/N and so these definitions were held as authoritative.

In every case of labelling, opponents were able to apply Wikipedia's policies as well as the traditional standards for reliable sourcing and citation established in academic fields to the case at hand. In every case, proponents failed to establish alternative standards by which sources from Wikipedians themselves might be regarded as reliable within Wikipedia's foundational policies. During the Wikimedia-l discussion, for example, Ziko van Dijk was able to apply established standards of designating primary and secondary sources and the reliability of each from the viewpoint of academic historians. Although proponents of oral citations called for an alternative to traditional encyclopaedic rules and the possibility of doing things differently from past encyclopaedias, they failed to appropriately translate Wikipedia's foundational rules and policies into an alternative standard. Without such a standard, filtered through the lens of policy, surr's traveling companions retained the labelling that opponents applied to it.

When controversy struck the sources used by editors in constructing the Egyptian Revolution of 2011 article, successful editors were similarly able to label certain sources and editors as illegitimate. [.68], for example, catalysed the removal of the death table, a list of facts summarising deaths in each city during the protests, when s/he brought the 'no original research' policy to bear on the actions of editors. [.68] was successfully able

to label the editors' actions as 'original research' in terms of policy and therefore illegitimate when respondent editors such as [Wipsenade] acceded by declaring that the table 'was out until confirmed' by independent sources. Editors could just as easily had referred to the exclusion within the 'no original research' policy for 'routine calculations'[54] but without the knowledge of such exceptions (or the desire to continue arguing with her/him), [.68]'s definition of the facts held.

On some occasions, experienced editors faced off against one another, both parties employing their knowledge of policies in order to advance the representation (or removal) of facts in the article. When [Ocaasi], for example, added Al Jazeera's live YouTube feed to the external links section of the article at the height of protests, [Cptnono] opposed this addition by applying policies, including 'Wikipedia is not a newspaper' (WP:NOTNP) and 'recentism' (WP:RECENTISM), albeit without including their URLs by reference. When [Ocaasi] responded by urging editors to 'Ignore all rules' (WP:IAR), [Cptnono] threatened him with disciplinary sanctions, and when one of [Ocaasi]'s allies agreed with [Cptnono], [Ocaasi] relinquished. The network acting against [Ocaasi] proved too strong for him to continue to oppose the link's removal.

The Egyptian Revolution article was able to travel further and with greater integrity than surr because it was also accompanied by editors who were able (and willing) to encode the article with statements in Wikidata. Such statements made the Egyptian Revolution facts travel both within Wikimedia sites and outside of Wikimedia to sites employing Wikidata to represent facts from Wikipedia. Surr, on the other hand, contained no Wikidata statements and was only represented in Wikidata as an empty

---

[54] See https://en.wikipedia.org/wiki/Wikipedia:No_original_research#Routine_calculations.

number with links to three other articles at the time of this study. The Egyptian Revolution article also demonstrated how some of the fact's traveling companions remained invisible to the reader but were critical to the facts' travel. Google, social media sources and Wikipedians all played an important role in ensuring the travel and support of facts, but remained hidden to the reader because their visibility would bring into question the boundaries between Wikipedians as passive collators of all human knowledge and sources of information that are tainted by ideological passions and biases.

The examples of both surr and the Egyptian Revolution of 2011 indicate how important facts' traveling companions were to the travel of facts embedded within them. Traveling companions included Wikipedia policy objects, Wikipedia editors and the actors represented in sources used to support facts contained within the article. Traveling companions were able to successfully support the travel of facts when they were deemed legitimate by other editors (in the case of sources) or when editors recognised the authority of editors who supported a fact's travel. Such authority was, in turn, determined by the ability of editors to deploy appropriate policy objects in order to either prevent further debate, or to convince editors that their interpretation of policy was superior and thus enlist other allies that would then prevent further debate.

7.1.2 Terrain

Wikipedia's coded infrastructure played an important role in defining the ideal environment for facts and therefore had a significant impact on how facts travelled within it. Infrastructure, as constituted by policy and technological tools, determines what constitutes 'good' traveling companions, what is 'suitable' terrain and which are 'characterful' or 'replicable' facts. Whereas policy (as outlined in chapter four) defines at

a foundational level what knowledge consists of and who is best placed to represent that knowledge, Wikipedia's coded tools serve to reinforce particular interpretations of policy to advance certain ideologies above others.

In the case of surr, the terrain in which facts travelled proved to be incredibly rocky and difficult. The article and its sources faced significant challenges by opponents along the way. The lack of sources in the standard format required for articles meant that there were few, if any sources, available for editors to choose from. Furthermore, the consequences of Wikipedia's verifiability logic created restrictions that meant that surr, as a fact that is held in the minds of its players rather than in multiple sources in books, journals and media articles, was of a particular disadvantage.

There is still very little published about the sport of surr in material form. The only citation from the English Wikipedia version still available online is to the jagran.com article indicating that surr games are becoming extinct in the villages of northern India. This does not mean that there isn't a rich body of knowledge about surr among the people who have experienced it. Because knowledge-holders have not translated their knowledge into the appropriate material format required by Wikipedia English editors, however, their knowledge remains unacknowledged, their voices unheard by readers of the surr article in English. The inability of the publishing market to supply such materials therefore has a significant effect on how far surr could travel since Wikipedia articles are so dependent on the sources from which they derive or support their facts.

Furthermore, the ideological strains inherent in Wikipedia knowledge work had a detrimental effect on the travel of surr because the oral citations that supported its facts challenged the principle of verifiability. Verifiability, as articulated by policies discussed

in chapter four, is a powerful foundational logic and ideology underpinning all Wikipedia work. The principle of verifiability is conceptualised within policy as enabling a communicative act in which the reader is able to request a source for an unsupported fact and to follow the citation in order to verify whether the source has been accurately reflected in the article. Verifiability is not just a standard by which authorship is attributed on Wikipedia; it is foundational because it articulates the kinds of rights and freedoms enabled by its practice. Although citations could be followed by the reader, they were not deemed to have been collected by an independent source and so were labelled as unreliable to English Wikipedia.

Verifiability enables communicative action and thus participation by editors that can be framed as emancipatory, but verifiability also limits what facts can be added or stabilized within Wikipedia to those that can be supported by sources that fit a particular material form. The affordances of the internet technically enables the expertise of any one of its users to be recognised but, in this case, the power embedded in the verifiability discourse proved too strong for proponents to dislodge and so their expertise remained unacknowledged.

Analysing the trajectory of the Egyptian Revolution of 2011 article, facts had a much smoother ride. There was a healthy market of sources available, many of which were packaged in the standard format indicated by policy and practice. Although some facts were challenged by editors, their removal did not compromise the entire article because it was already supported by numerous other facts deemed acceptable by editors.

The predominant sources employed to support facts in the Egyptian Revolution article were not considered as the most reliable according to policy, but the sources enabled

verification by readers and were therefore largely unchallenged. Where they were challenged were cases in which editors believed that the sources were not being accurately reflected in the content of the article, such as the case of inaccurate figures in the death tables according to [.68].

Academic sources are positioned in terms of policy and discourse as the ideal, but since those sources are not always available for the subjects that Wikipedians are covering, other sources are used in practice. Editors still advance a hierarchy in terms of which types of sources are most reliable in context, but a different type of source was then held up as the ideal. In the case of the Egyptian Revolution article, this ideal was encapsulated in the Al Jazeera source that was deemed to most accurately represent the subject of the protests.

This doesn't mean that academic sources are no longer considered ideal by editors. Academic sources are positioned as ideal, but when enough people are motivated to write about a subject on Wikipedia, they will find a way of interpreting policy in order to meet their needs. This is yet another example of how Wikipedians interpret policy according to their needs in the context of the article. It also means that, if editors want to write about breaking news, they will have to keep making judgement calls about which sources are most reliable, given the general dearth of academic sources about those topics, as well as the fact that the complexity and depth of academic sources, even if relevant, doesn't suit the rapid fact-adding practice of breaking news accounts.

The vocabulary and communicative actions enabled by Wikipedia's verifiability logic influenced the terrain in which facts travelled in both cases. Verifiability was recognised by editors as even more important as achieving the sum of all human knowledge

because it was the foundation of Wikipedia's claims to authority. There is, however, an inherent bias to a stringent application of the verifiability principle that has an impact on what Wikipedia can represent. While traditional encyclopaedias such as Encyclopaedia Britannica are their requirement for subject matter experts to write about articles and are thus limited by which experts believe is important, Wikipedia is similarly limited by what the experts that *it* claims are legitimate believe is important. Although Wikipedia extends this notion to a wider range of sources including journalistic sources, there is still a bias towards sources that are online, accessible, immediately available and that can support separate claims within an article, rather than those that provide wider analysis.

7.1.3   Character

A fact's traveling companions and terrain influence to a great extent how far and well it travels, but the character of facts themselves is also instrumental to the success or failure of a fact's traveling well. In the context of Wikipedia English, where the majority of editors are from the United States and Western Europe, a fact's character is determined by what editors find interesting, as well as how replicable a fact is within Wikipedia's socio-technical system.

In the case of the Egyptian Revolution, the massive scale of media attention by the English-language press, particularly in the United States, indicated that events in Egypt in 2011 were highly interesting and surprising to a wide-range of audiences in Western countries. Surr, on the other hand, remains very thinly described by the media or academia, indicating a lack of interest by local as well as international knowledge inscribers.

Every representation on Wikipedia is the result of a series of collaborative actions. Facts travel well when they are accompanied by traveling companions who are deemed suitable, within a terrain that is conducive to their travel, and when the facts themselves are interesting to relevant users. Wikipedia's traveling facts were demonstrated to travel well when they constituted alliances between human and non-human actors. Allies in Wikipedia's context are actors including individuals (editors, administrators, users), institutions (for example, the Wikimedia Foundation), texts (other Wikipedia articles/language versions/project pages and sources) that strengthen a human actor's case in achieving their goals. A key goal of most Wikipedia editors is to have their edits prevail, and edits are a representation of an editor's ideological perspective.

In the next section I outline the findings relating to how authority is configured within Wikipedia and whether such configurations constitute a reconfiguration of authority and expertise in relating to previous encyclopaedias.

7.2   The reconfiguration of authority and expertise

This thesis demonstrates that there are three key ways in which authority and expertise are being reconfigured within Wikipedia. Investigating who are considered authorities by Wikipedia editors in terms of both policy and practice, it was revealed that certain sources and editors are considered more authoritative than others within Wikipedia's socio-technical environment, and that the identities of editors and sources is a significant part of how Wikipedia maintains its authority as a source of facts about the world.

7.2.1   Sources' authority

Authoritative sources in Wikipedia are those that are defined as such by editors during discussions and the sources that remain after editors have stabilized an article. Authority was enacted by editors such as [Wipsenade], in the case of the Egyptian Revolution article, who stated that the 'BBC, UN, EuroNews, CNN, NHK, Reuters' were the sources that would be able to confirm figures in the death table. Discussions among editors about which sources are most relevant in the context of the article not only serve to identify which sources are considered reliable by editors at the time; such discussions also serve as pointers to editors about what will be more likely to be accepted when they add sources to particular claims in the future. In this way, rhetoric by individual editors in conversation with one another serves to reinforce which sources are considered reliable by their constant enacting. It is not enough to *call* a source 'reliable' or 'authoritative' during discussions, editors must also *enact* this authority by adding particular sources and making sure sources remain linked to facts in the context of regular editing over time.

In the case of surr, opponents were able to successfully label sources as unreliable, primary and non-compliant in terms of Wikipedia's standard copyright agreements. The remaining sources in the article at the time of this study included an article in the online newspaper Jagran.com and a citation to a weblog that had been taken offline. Despite the fact that the citation to the blog post was dead, it remains on the article because proponents recognize that even empty references to material, independent sources, even if they are to blog posts that are not ideal according to policy, are better than fewer citations. The fact that the dead citation has not been removed indicates its importance in an article with very few citations when the authority of Wikipedia articles is signaled by the appearance (rather than the functioning) of citations.

Analysis of sources deployed by editors over time in the case of the Egyptian Revolution of 2011 article, as well as those that remained after the article stabilized indicated the dominance of certain media sources over others. Sources, most notably Al Jazeera which also dominated the media environment in the Western world during the protests, were found to be influential allies in the context of the article. Social media sources, on the other hand, were progressively removed from the article over time.

There is evidence that the authority of sources on Wikipedia is being reconfigured in terms of earlier encyclopaedias. Sources of knowledge are defined more widely in Wikipedia practice than they are in Britannica, for example. Wikipedia policy and practice indicates that sources can be to a wide range of authors and publishers, from both institutional and individual self-published sources, and from traditional media organization to weblogs and social media sources, even if, in practice, some of these sources are removed over time. The authority of an author in a particular topic is not only a product of an in-depth understanding of a phenomenon on Wikipedia, as it was in the case in traditional encyclopaedias. The authority of sources is, instead, a product of their materiality on Wikipedia, where a source's accessibility makes it most likely to be employed, rather than the author's expertise in the subject.

In the case of breaking news articles such as the Egyptian Revolution article, accessibility was determined by the accessible style (particularly in the case of academic articles), the availability of a source immediately after events had taken place as well as the modularity of sources. Accessibility bias on Wikipedia means that the more rapidly available a source is, the more likely it is that editors will choose it (as in the case of Al Jazeera), and that short articles supporting particular statements rather than in-depth monographs or longer academic articles are more likely to be employed as sources.

Whereas for Encyclopaedia Britannica, the authority of the source of one of its articles is the author's long-term and deep engagement with a particular subject, on Wikipedia authority is the product of the material features of a source that make is useful under the conditions of editing, especially in the case of breaking news articles. Although Wikipedia policy indicates that sources need to be reliable and that reliability is indicated by sources' publication process or by authors who are regarded as authoritative, the sources that are actually used on Wikipedia are sources that represent the best (as determined by the policy on reliable sources) of what is available and accessible at the time they are required. The sources that are then used and remain unchallenged in an articular act as future evidence of a source's reliability, even if the choice was initially determined as a result of contextual factors that are no longer relevant.

7.2.2   Editors' authority

Wikipedia has mechanisms for ensuring that editors are disciplined by others who have authority by virtue of their power to delete articles, ban users and protect pages, among other actions. There is, however, no editorial oversight for the project; no core group of editors who decide what the project should cover at the outset and who can deploy resources in order to ensure consistent coverage of such topics as in traditional publishing projects. Instead, authors write articles about topics that they are interested in. When editors have different ideas about what should be included or excluded, they attempt to exert their authority on other editors through multiple means.

Means available to editors includes both persuasion and direct action on the article itself. Editors can advance their solution or perspectives on the talk pages or in forums

265

using the vocabulary and discourses of Wikipedia policies and principles. Such persuasive means were demonstrated in the case studies by editors like [Fifelfoo] who, in the case of surr, was able to successfully label the oral citations as unreliable and by [Cptnono] who was able to lobby for the removal of the live Al Jazeera link as unencyclopaedic on the relevant talk pages.

Authoritative editors are also those whose direct action on the article itself is acceded to by other editors. Such editors work to destabilize or stabilize the article, depending on their ideological perspectives. In both case studies, the ability of editors to classify their actions as complying with encyclopaedism by deploying a variety of objects (including policy objects, warning tags, categories and citations) enabled them to prevail. Policy objects, for example, acted as frames that could be used by both the proponents and opponents of the facts under discussion in order to label sources by applying different aspects of Wikipedia's policy vocabulary. In the case of surr, opponents of oral citations deployed the vocabulary from Wikipedia's verifiability policy regarding what constituted reliable sources to declare that the oral citations were unreliable. Warning tags and other objects were used to destabilize the article by discrediting it according to what [Fifelfoo] and others saw as non-compliance with Wikipedia rules.

Since articles are never complete, the goal can never be to make the final edit or to have the last word before the article goes to print as it might be in traditional formats. Instead, successful editors can only make iterative movements towards either stabilizing the article so that it becomes more difficult to attack it, or destabilizing the article by removing its allies and otherwise discrediting it.

Adding text, categories and links, for example, are defensive moves to pre-emptively defend the article against attacks on the grounds of non-notability, lack of categories or sources. Similarly, destabilization can be achieved by those who disagree with either the article or the authors of the article by proposing articles for deletion, reverting the edits of others, adding warning or maintenance tags, or disputing the validity and relevance of the topic. Strategies for stabilization and destabilization include the ability to summon or to cut off access to allies (in the form of actors and other texts) and to control the visibility of coded objects.

Table 7.1 illustrates the actions taken by proponents and opponents of articles and the facts contained within those articles in order to either stabilize or destabilize them. Actions were directed at three lines of attack: the article level, actions taken at the level of the article itself, and actions taken against (and in support of) editors themselves. Actions that include persuasive means are in black while direct actions by editors at the level of the article itself are in red.

Table 7.1 Strategies for stabilizing and destabilizing articles by editors

Lines of attack	Stabilization actions	Destabilization actions
Article level	Defend article against deletion; Ensure article does not raise red flags in New Page Patrol by adding categories, sources and interwiki links	Nominate article for deletion; Remove coded objects such as categories and sources so that the article will become visible in watch lists and working groups
Within the article	Remove warning tags; Revert rivals' edits; Add content/Delete rivals' content; Add coded objects; Defend sources as reliable; Defend article as notable.	Add warning tags; Revert rivals' edits; Delete rivals' content; Remove coded objects; Define sources as unreliable; Define article as not notable.
Editors	Report rival editors to higher authorities/Defend allies during disciplinary procedures; Praise the edits of allies and criticize the edits of rival editors.	Report rival editors to higher authorities/Defend allies during disciplinary procedures; Praise the edits of allies and criticize the edits of rival editors.

Note: Text in red indicates direct action taken to the content of the article, black text indicates persuasive action taking place within forums.

Source: author

The difference between these two types of communicative action is important to note. Editors will not be considered authoritative in the context of Wikipedia unless they participate in direct action. Edit count is the most important heuristic used by editors in order to evaluate authority on Wikipedia (see Jemielniak, 2014) and, although edits to a talk page are also considered in the edit count, edits to the article itself are a quicker, more direct way of ensuring that facts are represented according to the ideologies of participating editors.

The principle, Wikipedia:Be Bold suggests that editors act before waiting for consensus since consensus through discursive mechanisms on the talk pages and other discussion forums is slow and relatively costly to attain.

> **Wikipedia:Be bold** (WP:BOLD) can be explained in three words: "Go for it". The Wikipedia community encourages users to **be bold when updating the encyclopedia**. Wikis like ours develop faster when everybody helps to fix problems, correct grammar, add facts, make sure wording is accurate, etc. (Wikipedia:Be_bold)

Editors are encouraged to edit the article directly rather than to wait for consensus on every action but this also encourages counter actions by rival editors who do not agree with the ideology being advanced by particular editors. Thus, articles tend to progress through the stabilization and destabilization of articles, which constitutes a form of dialogue among editors and a notification mechanism to readers. Adding a neutrality warning to the article, for example, not only constitutes a communication action from an editor to readers about the content in the article; it is also a critique by the editor that s/he believes the articles authors are not appropriately applying Wikipedia's core NPOV policy.

There are numerous fronts that rivals can attack and defend from: at the article level, within the article and among editors themselves. Each has benefits and challenges. Attacking the article when it is first published can be the most efficient method of attack because there are fewer stakeholders who have become involved in the article and therefore fewer rivals to defend against after the attack, and a high degree of interpretive flexibility with which to attack the article since it can be interpreted in numerous ways. On the other hand, such attacks can be difficult if there is momentum

behind the article, and where numerous sources exist to support its relevance to the rest of the corpus.

Examples of communicative action at the article level included reverting, deleting, removing citations and adding coded objects. The most fundamental disruptive action available to editors is the revert. Any editor is able to revert the article back to the version before the rival editor had made their changes. The revert effectively deletes the rival editor's changes and disregards their labour. The revert is at the heart of edit wars, events where editors continue to revert one another's edits in a process that sets the article in continual back and forth motion, causing instability and disruption that can only be halted by administrative remedies or locking the page to edits by particular individuals or editor groups. Reversion is used to remove vandalism because it is a quick and easy mechanism to remove disruptive content.

Another disruptive action available to editors is for them to delete another person's edit. The editor will delete the element (text, image reference, citation, category etc.) and optionally use the edit comment to explain their deletion. Although this comment is visible to all and has a purpose in summarising edits for all editors, the primary communication is between the editor and their rival. The rival can then choose to revert the edit back to their version or to concede.

Editors can also remove citations and sources because they believe too many sources are being used so that the article is loading too slowly, because they do not believe the source to be reliable, or because the accompanying text is not accurately reflected by the source. Editors may also remove citations of rival editors on the pretence of these reasons but fundamentally to disrupt the rival editor's work.

Populating the article with content is one of the most important ways to stabilize the article. Adding citations to sources that are widely regarded by editors as reliable (that is, used without question elsewhere) as well as adding assets such as images, graphs and infoboxes, all add to the stability of the article. Wikipedia's coded objects constitute allies that need to be attacked by rivals, thus making the task more difficult to accomplish.

When policy objects are brought by both proponents and opponents, the victorious party is the one who is able to convince others of their definitions or labels. The evidence of editors being convinced usually does not take the form of acquiescence in conversations, but rather the cessation of debate and the lack of further challenges to the article. Consensus, then, does not mean agreement on Wikipedia. It means that, for various reasons, editors have stopped arguing.

One of the reasons for such a cessation could be that editors have come to agree with their opponents. In the case of both surr and the Egyptian Revolution, however, many of the parties who were not successful in representing their views or facts stopped participating in debates because they lost hope and surrendered. Achal Prabhala, for example, stopped participating in debates because he became resigned to how he believes Wikipedia continues to uphold the authority of knowledge in the printed word. In the case of the Egyptian Revolution article, editors attempting to have a reasoned dialogue about whether the article title should be changed from 'protests' to 'Revolution' surrendered to the will of the mob that continued to change the title before agreement had been reached. The fact that the article's title changed before the final verdict on the discussion had been made is testament to this act of surrender.

What Wikipedia calls consensus might, then, be more accurately called 'acquiescence', signified by passive assent or submission, rather than active agreement by editors and other stakeholders to the more authoritative actions of their rivals. Importantly, authority in Wikipedia is signified by the greater ability of certain editors to translate their ideological perspectives using the vocabulary and tools available to them. Since the vocabulary and tools limit what can be represented on Wikipedia, authority is necessarily limited to those whose perspective concords with the values and principles embedded within Wikipedia's coded environment.

7.2.3 Wikipedia's authority

Authority is also being asserted in the advancement of Wikipedia, itself, as an authoritative source of facts about the world. Whereas Britannica asserts its authority by virtue of its ability to discover and contract expert authors to write encyclopaedic articles, Wikipedia's authority is derived from the fact that authors do not apply their own expertise to a particular topic, but instead apply the knowledge represented by reliable sources to each claim that is made.

The authority of Wikipedia is dependent on editors retaining this role as passive collectors rather than active creators of knowledge. Since all knowledge construction requires at least some level of editorial decision-making by authors, however, only certain editors who are able to translate their knowledge into the norms of Wikipedia, are able to have their edits maintained. Editors who are able to find sources to support their knowledge of particular facts or to apply policy to the representation of other facts, are able to apply their own knowledge to the construction of Wikipedia articles. Those

who fail to translate their knowledge of facts according to Wikipedia's policies and norms, as happened in the surr article, are consequently discredited and expelled.

Furthermore, Wikipedia is able to reinforce its authority through its identification with the libertarian ideology encompassed by the verifiability principle. According to Wikipedia's verifiability rules, readers have the ability of readers to question the authority of particular claims contained within the article and citations to traceable sources in articles reinforce the authority of claims contained within articles. The assumption by readers (one that is evidenced by studies of Wikipedia's credibility mentioned earlier) is that the appearance of sources in an article is a *sign* of credibility but that readers seldom investigate themselves.

Authority consequently becomes a key strategy used to convince the world that Wikipedia is reliable and a way of stabilizing articles along particular ideological lines. In other words, the same strategy occurs both at the macro and micro levels of Wikipedia practice, but in different instantiations. At both levels, there are unexpected consequences to Wikipedia's claim to authority. Wikipedia's authoritativeness persuades readers of the credibility of all content by virtue of the appearance of sources and this can lead to a lack of critical investigation who fail to investigate whether the article has been accurately represented, and whether the sources are, in fact, reliable. Wikipedians' authority, in other words, often prevents further debate and criticality from readers.

7.3   The distribution and sources of power on Wikipedia

Analysing the Wikimedia Foundation's organizational chart or the technical permission levels of Wikipedians, power on Wikipedia would appear to be hierarchical. The

Wikimedia Board of Directors are considered to be at the top of the hierarchy, the Wikimedia Foundation below them, followed by Wikipedia and other Wikimedia projects in all available languages, and finally the ordinary users of particular projects at the bottom of the pyramid. Wikimedia projects are, however, unlike traditional organizations in which rules are constructed and applied by a central managerial force. Wikipedia is also unlike a traditional encyclopaedia in which editorial decisions are made by a core editorial team with the work of writing articles delegated to particular authors. Instead, Wikipedia's editorial decisions are guided by the actions of individuals at the edges of the network.

The Wikimedia Foundation (WMF) is charged with supporting the efforts of the global volunteer community. The majority of WMF employees support the development of the software and other technology that constitutes critical infrastructure for Wikipedia and other Wikimedia projects. There are strict rules about WMF employees editing content, particularly because the WMF is able to maintain immunity from prosecution by virtue of its identity as a publisher rather than author of content that it hosts.

The board of the WMF has ultimate authority over the affairs of the Foundation and regularly makes resolutions that are binding on all Wikimedia projects. The WMF board tends not, however, to make editorial decisions about projects and focuses, instead, on the approval of WMF finances and strategic plans, as well as resolutions on amendments to the terms of use of Wikimedia websites and other legal issues.[55] When the WMF board has, on rare occasions, attempted to make decisions that impact on the practice of volunteer editors, they have faced major pushback from the community.

---

[55] See https://wikimediafoundation.org/wiki/Resolutions.

In 2011, for example, the WMF board made a resolution about 'controversial content' on the site that was catalyzed by the appearance of a vulva on German Wikipedia as the 'article of the day'. The board resolution urged the Wikimedia Commons community to more actively curate content and remove media that did 'not meet existing policies and guidelines for inclusion'[56] but also asked the Executive Director of the WMF to 'implement a personal image hiding feature that will enable readers to easily hide images hosted on the projects that they (did) not wish to view'. Two months later, after rigorous debate and protest, particularly by German Wikipedians, the board rescinded its request for an image hiding feature. Although a movement-wide referendum (with a record of 24,000 editors participating) indicated a majority of editors in favour of the filter, a German Wikipedians voted 86% against the feature and there were serious discussions about whether German Wikipedia would fork if the Foundation imposed the feature on German Wikipedia.[57] Just because an actor within Wikimedia's decentralised network has the power to make a decision, doesn't mean that they will make such decisions.

Similarly, although the power and equality of every Wikipedia project (consisting of 290 language versions) is often asserted, in reality, the majority of Wikipedia projects are heavily dependent on the larger Wikipedias to guide decision making. English Wikipedia often dominates discussions, since discussions on mailing lists tend to be dominated by

---

[56] See https://wikimediafoundation.org/wiki/Resolution:Controversial_content.

[57] See https://meta.wikimedia.org/wiki/Image_filter_referendum/Sue%27s_report_to_the_board/en.

English Wikipedia editors and since English Wikipedia attracts significantly more readers, even in countries where English is not the official language.[58]

Central to Wikipedia's core goal is that it is a project to build free encyclopaedias in all languages of the world. Wikipedia co-founder, Jimmy Wales, has urged editors to write articles in their own language from scratch, rather than only through translation since translation effectively implies the domination of one culture over another. Wikipedia is not neatly separable into different language versions that are free to represent their own perspectives on topics, however. The diversity of Wikipedia's language versions is limited by a number of different factors including the fact that editors aren't required to speak the language of the Wikipedia project and can make edits based on limited understanding of the language. Many editors work on multiple language versions, bots work on multiple language versions and core policies are shared among all versions.

Instead of centralized editorial control, then, decision-making on Wikipedia happens at the edges of the network, with editors directly enacting their perspectives within the content of the article and/or discussing, announcing or persuading others of their changes in related discussion spaces. Decentralisation does not mean that there are no power imbalances or hierarchies but that hierarchies develop at the edges of the network in ways that are dynamic rather than static. Editors regularly come and go from the network. Experienced editors align themselves with more experienced editors and authoritative allies from stakeholder groups such as the Wikimedia Foundation; they also deploy objects in order to shore up support or destabilize their opponents. Inexperienced editors participate without translating their perspectives using coded

[58] See https://stats.wikimedia.org/EN/Sitemap.htm.

objects or policy objects and therefore have little sustained control over what is represented.

Applying the theory of boundary work, boundaries over what Wikipedia represents are being established by ad hoc actor networks, driven by editors with different motivations and concerns and equipped with a variety of tools and editorial objects. Instead of decision-making according to rational discussion among actors who have an equal ability to impact the decision-making process, decision-making occurs through a process of persuasion and direct action on the article itself in which participants have widely divergent knowledge of Wikipedia's vocabulary. Decisions are effectively made, not by all of Wikipedia's editors and stakeholders, but by local groups, where the authority of editors is determined by their ability to use the tools to their advantage rather than by virtue of their knowledge or expertise in a particular topic.

Wikipedia is a network that extends from the facts that it represents within articles to the objects (including policies and tools) and processes that become enrolled as the fact travels along the network. Instead of each language version of Wikipedia governed by separate language committees and national chapters, and finally governed by the WMF and the WMF board, power is dispersed throughout the network in various assemblages of actors, objects, practices, relationships and processes.

This decentralised view of power within Wikipedia has significant implications for any actor, including the Wikimedia Foundation itself, to be able to influence the direction of the encyclopaedia, particularly in view of the heavy skew in terms of demographics and the wide gaps in coverage of global topics. No single actor has the power to change the corpus as a whole. The Wikimedia Foundation cannot mandate editors to change

direction and focus on particular areas, or filling particular gaps. Change can only occur through the small actions that perform Wikipedia as an encyclopaedia that covers certain subjects, and affords for certain types of actions. Each decision in the form of an edit is a speech act that is performing a particular view of the world, a particular definition of what Wikipedia is, what the role of Wikipedians is, what sources Wikipedia should use.

Consensus by the entire organisation, editing community and various stakeholders is expensive, time-consuming and difficult. As such, it happens in only few occasions. Instead, successful expansion of the encyclopaedia's goals and practices can only be obtained through iterative performance of a particular encyclopaedic identity and through such actions editors can demonstrate that, since Wikipedia already occupies such territories in some subjects, there is a case for expanding such territories in other areas.

In the case of the oral citations debate, for example, Achal Prabhala chose the movement-wide mailing list to discuss the oral citations, instead of fighting the battle directly on the reliable sources noticeboard and the talk pages of the articles. As a result, the opposition to the oral citation project defeated their proponents, using discursive techniques and speech acts in order to destabilize the citations by cutting off their allies. Editors didn't only discuss whether the content was appropriate or inappropriate as a way of defining what Wikipedia should include or exclude; they actively performed what should be included or excluded by shaping the articles themselves. Consequently, those who are more effectively able to perform according to Wikipedia's embedded logics and vocabulary are more successful than those who merely possess expertise in a particular subject.

278

In the Egyptian Revolution article, editors were actively performing a boundary between what they considered authoritative and reliable sources for the Egyptian Revolution by adding certain sources to certain parts of the article and deleting others over time. The pattern that emerged of which sources were sustained over time was not only a reflection of the ideological interests of its editors; such patterns serve to reinforce the authority of such sources more generally.

In both cases, the power of technocratic editors had an impact on the representation of the subjects of the articles, and on the ways in which similar subjects will be represented in the future. Each successful decision that is made by actors becomes part of the stock of evidence used by editors for future decision-making. In the case of surr, decisions by opponents to the oral citations makes it more difficult for the use of oral citations on the encyclopaedia in the future; the success of the Egyptian Revolution article increases the chances of success by editors wanting to write about breaking news topics. In each case, successful editors determined the boundaries of what Wikipedia will represent and how it will represent those subjects for now and the future.

It is important to note that these successful moves don't mean that success is being determined once and for all. Success is always temporary. Boundaries are being continually constructed, and although they can become congealed through material objects over time and thus more difficult to dislodge, there is always the potential for new combinations of actors to dislodge that power over time.

7.4 Conclusion

At the beginning of this thesis, I outlined the co-production idiom as a useful way of framing the effects of socio-technical change on the material landscape. Jasanoff (2006) writes that, during time of significant social and technological change, traditional authorities are often reaffirmed, while at the same time there is a redefinition of identities as societies redraft social order in order to fill gaps in expertise that result from the adoption of new technologies. New power relations become lodged within identities, institutions, languages and representations.

The findings of this thesis indicate that, within Wikipedia's network, power is lodged in the identity of the authoritative source and the experienced Wikipedian. Power is also lodged in discourses that frame knowledge as that which is represented by a passive, independent analyst, and in representations that serve to either reject or remove the voice of the subject over time as social media sources and oral citations are rejected. Technological and social change, in other words, have not only involved the introduction of new roles, predominantly in the form of the experienced Wikipedian, but also the re-affirmation of existing authorities, notably the academic expert and the institutional professional. The Wikipedian has become another 'quintessential bridging figure of modernity', arising out of the need to translate the growing complexities of science and technology for broader social goals (Jasanoff, 2004, p. 39). In Wikipedia's case, such goals have focused on access by people around the world to the sum of all human knowledge – but with knowledge defined by editors as that which is traceable and verifiable according to particular standards.

As Jasanoff notes, in order to accommodate significant change there tends to be a re-drafting of the rules of social order regarding trustworthiness and authority. This

redrafting is neither clear-cut nor does it happen immediately. Instead, boundaries between expert and non-expert, scientist and non-scientist are being continually reinscribed. In some cases the authority of certain experts is reaffirmed; in others, there is a translation of one type of expertise into the context of another as new expert identities are introduced.

In the case of Wikipedia, the authority of academic professional has been reaffirmed according to policy and practice which often positions such sources as the ideal. Sources have also extended beyond academic professionals to a variety of emerging authorities including journalistic sources, although power tends to concentrate in a few central English-language sources at the expense of lesser known local sources. Such redrafting is neither clear-cut nor does it happen immediately – only time (and further research) will tell whether patterns exhibited in these cases will be sustained over time.

Power within Wikipedia is the result of the co-production of technology and society. Authority cannot be achieved through the actions of Wikipedians alone but rather by its readers' legitimation of Wikipedia's power to represent. Technical permissions enable certain editors greater power over the editing process than others, but social norms dictate whether it is appropriate for editors to use this power to determine editorial outcomes. Authority on Wikipedia and of Wikipedia is co-produced, with knowledges, expertise, technical practices and material objects 'shaping, sustaining, subverting or transforming relations of authority' (Jasanoff, 2006, p. 4)

Instead of being populated by amateurs, Wikipedia has become a site in which highly skilled individuals are working to represent human knowledge according to a particular ideological frame. Many of these individuals are driven by Wikipedia's goal to represent

the 'sum of all human knowledge' but there are consequences to the ways in which Wikipedia works and what its epistemological foundations are to what facts are best represented within it. Wikipedia does not negate the need for gatekeepers as was earlier claimed by commentators such as Yochai Benkler; Wikipedia has demonstrated that gatekeepers are a necessary feature of information environments, particular in the realm of facts. As Wikipedians working on the edges of the encyclopaedia decide what it should cover in the future, they are determining not only the future of Wikipedia, but the future of others who are in the business of creating and representing facts.

Chapter 8: Conclusion

As Jasanoff writes, the ways in which knowledge is taken up within societies have an impact on who is considered the expert, subject and revolutionary. This thesis has demonstrated how Wikipedia has become an authoritative source of facts about the world. Authority on Wikipedia is enacted by the reaffirmation of the ideal knower as passive, independent and removed from the experience of a phenomenon, but in doing so the subject of the article is relegated to a voiceless position. Some of those whose experience of the world is not already translated into a material form, and have failed to translate their ideological perspectives using the rules and tools afforded by Wikipedia's environment have become revolutionaries against what they see as conservatism or Western ideological bias.

Information from Wikipedia articles is now used as the answer to questions about people, places and things in search engines and as a foundation for both artificial intelligence agents and other sites employing Wikipedia data. How had Wikipedia attained this authority from the days in which it was either derided as a poor quality alternative to traditionally published encyclopaedias, or lauded as a significant alternative to traditionally published encyclopaedias, but never as the dominant producer of facts, at least on the English-speaking internet?

Wikipedia has developed around the logic of verifiability and it is this logic that influences the way in which representations are constructed on the encyclopaedia. Analysing how two sets of facts that lay on the boundaries of what is considered acceptable for inclusion on Wikipedia were debated and acted upon, We have seen in

the preceding chapters how the power to represent facts is being reconfigured by Wikipedia in the context of larger socio-technical forces.

The identity of the expert, the subject and the revolutionary, have been reconfigured as a result of the ways in which knowledge is taken up, as Jasanoff (2004) predicts. Far from original speculations that Wikipedia would reduce the importance of academic authority, the authority of the scientist or tenured academic has been reaffirmed by Wikipedia. This reaffirmation has been enacted by policies such as the reliable sources policy and in the practice of denigrating certain sources by comparing them against the academic ideal.

Other sources of knowledge from media organisations, civil society organisations and government, however, are also accepted as legitimate sources in practice. By publicly reaffirming the identity of traditional sources of factual knowledge in the context of academia, and by deploying alternative sources in cases where academic sources are unavailable, Wikipedia is reconfiguring which sources of knowledge obtain more attention and consequently, which sources become more authoritative.

In parallel to both the reconfiguration of the authorities of knowledge sources outside of Wikipedia, Wikipedia has also become the site for the rise of a new class of experts. Successful deployment of sources is the product of a Wikipedia editor knowing what will be considered authoritative in terms of norms, policy and the context of the article's subject area, as well as how to defend particular sources using the vocabulary particular to Wikipedia work. Only certain Wikipedia editors exhibit the expertise necessary to stabilize a particular assemblage of facts and sources in an article, and thereby impose their own perspective on the phenomenon.

Wikipedia's experts attain authority, not only by describing their criticisms or praise of facts through discussion on talk pages and relevant discussion groups, but by enacting their ideological viewpoint in edits to the article itself. Although still important, especially in terms of establishing precedents for future evaluation, an editors' discussions about what they believe is right or wrong about the article in terms of policy and norms must be accompanied by direct action to the article itself. Authoritative editors are those who are able to either stabilize or destabilize an article in concert with their ideological interests and strains. Since force of numbers is also important to the success of particular facts over time, alliance building is necessary in order for others to help provide the necessary support that will maintain this stability or instability over time. What is achieved out of stabilization has been called 'consensus' according to Wikipedia policy documents, but effectively stabilisation is the result achieved by the cessation of debate, not necessarily agreement by all stakeholders.

Wikipedia editors themselves, then, become a significant source of knowledge accompanying the travel of facts through Wikipedia's socio-technical system.

The work that editors do in framing articles, summarising sources, choosing topics, images and other assets is neither passive nor without perspective. Wikipedia editing, is, instead, a creative process in which one path among many is continually chosen. Those editors who are able to best perform the ideal identity of editors as passively collating and representing what already exists in independent sources are those who prevail. Their power is a product of their ability to make a complex system work for their goals.

The analysis of power and the powerful within Wikipedia's socio-technical system signifies the key contribution of this thesis to current debates and it is in this arena that lie fruitful grounds for further theoretical development.

8.1 Key contributions

Until now, theoretical frameworks with which to study Wikipedia have been concentrated on the removal of the figure of the gatekeeper. For Yochai Benkler (2006), for example, the figure of the gatekeeper in the traditional media organisation represented the most important power broker of the past. The gatekeeper stood at the gate to publication and was thus able to make decisions as to whether particular content and ideas were important enough to pass through. Because the gatekeeper was influenced by powerful corporate and government sources (particularly through advertising revenue and the availability of easy-to-reach state spokespersons), s/he tended to make biased decisions that allowed only certain subjects to be represented. According to this view, gatekeepers tended to represent the interests of the already powerful.

The participatory Web, according to scholars like Benkler, removed the figure of the gatekeeper since publishing was now cheap and easy and mediators were no longer necessary since people could publish on their own or in loosely connected groups. As noted in Chapter one, Benkler recognises the Web as liberatory.

> The various formats of the networked public sphere provide anyone with an outlet to speak, to inquire, to investigate, without need to access the resources of a major media organization. (Benkler, 2006, p. 11)

What we have witnessed in Wikipedia, however, is that new mediators have arisen out of the decentralised architecture of the internet and that these mediators determine what content is added and how that content is represented. Understanding exactly how this is occurring, as well as how the power of Wikipedians is mediated by the sources they represent, has resulted in a key gap in scholarship that this thesis has begun to fill.

Responses to Benkler (and others such as Weinberger, 2011) and their libertarian perspective in which the networked public sphere are viewed as enabling speech, have emerged from two camps: those who oppose Wikipedia's process and structure as liberatory, and those who provide empirical evidence of Wikipedia's hierarchies in terms of technical permissions (for example, Arazy, Nov & Ortega, 2014) and the increasingly fragile power of humans in relation to bots (for example, Niederer & van Dijk, 2010, and Geiger & Ribes, 2013). A more balanced approach that recognises both the agency of humans involved in the construction of technology and the diverse outcomes of such technologies is necessary, however, in order to understand *which* humans have more power than others in the context of an environment that is almost entirely mediated by code, but one in which power is not only a product of the ability to exert force.

Wikipedia is not governed by force; few editors are banned since banning is often ineffective when editors can use technical means in order to establish new accounts that evade disciplinary mechanisms. Instead, Wikipedia is governed by authority, a softer form of power where power is the product of an editor's knowledge of the environment, ability to speak the technical and legal language of Wikipedia, and to deploy objects and processes in ways that advance their ideological interests and are a result of ideological strains.

Although Gieryn (1983, 2001) applying Geertz (1973) provided a useful lens for understanding how ideological acts of boundary work take place when knowledge workers must identify what should be classified as in or out, two significant differences were discovered in the way in which boundary work takes places within Wikipedia's socio-technical structure. Firstly, since Wikipedia gains its authority not from the

authority of the editors themselves but rather through the authority of the sources that they use, rhetorical techniques are aimed at either flattering or denigrating the authors of the *source*, rather than other Wikipedia editors. Instead of Gieryn's scientists holding themselves up as ideal knowers by highlighting contrasts between rivals in a way that is flattering to themselves and excluding rivals by defining them as outsiders, Wikipedia editors hold up what they believe are the ideal sources in context and by defining rivals as outsiders.

Secondly, the practice of expelling and expanding sources within Wikipedia is not only a process of convincing others of the superiority of one's argument through the use literary devices of metaphor, hyperbole, irony and sarcasm, or syntactical devices of antithesis, inversion, and repetition, as Geertz argues. The practice of boundary work in Wikipedia is, instead, also a process of directly articulating one's ideological standpoint through editing the article itself. Instead of only talking about what is encyclopaedic vs. unencyclopaedic, Wikipedia workers must perform digital speech acts in order to advance or expel certain claims as facts.

Experienced editors argue their point or challenge other editors on the talk page or on mailing lists and are deemed authoritative by virtue of their ability to deploy terms from policy and practice, but experienced editors also take similar discursive action by editing the article itself. Such actions serve to either stabilize or destabilize the article, depending on the ideological position of editors. Editors add warning tags to disrupt the article, for example, or remove warning tags to stabilize the article. The vocabulary used to define the boundaries of the encyclopedia consists of both human-readable code and machine-readable code.

In order to advance this framework of heterogeneous communicative action, I combined Geertz's theory (and Gieryn's implementation) of ideological 'styles' together with Isin and Ruppert's (2015) conception of digital speech acts. This combination of theoretical approaches enables an accounting of the ways in which editors not only describe the boundaries between fact and fiction, or independent and advocatory approaches to editing, but also enact those boundaries through digital speech acts. Instead of digital speech acts being deployed only as a way of introducing rights claims as Isin and Ruppert articulate, however, Wikipedia's complex vocabulary introduces the idea that digital speech acts can also have the exclusionary effects.

The complex vocabulary deployed by experienced Wikipedians enables only certain actors to deploy objects and processes in order to stabilize or destabilize an article or convince others of one's claims. Editors must also possess the knowledge necessary to decode the traces left by other editors through their own editing practice. The effects of Wikipedia's complex vocabulary consequently serve to exclude those who do not possess the knowledge and experience necessary in order to participate in the representation of phenomena.

The exclusionary effects of digital speech acts can be applied to the other platforms that Isin and Ruppert mention in their positioning of digital speech acts acts as central to digital citizenship in order to understand power relations among actors within such platforms. The meaning of a *like* action on Facebook, for example, is determined through the affordances of the technology and the local norms attached to it; similarly, a retweet (RT) on Twitter can mean different things to different actors. A retweet can imply that the speaker is reasserting the claims being made by the original poster, for example, or it can mean that the speaker is merely exposing the original claim in an effort to

289

represent all views. The latter purpose is enacted when a journalist retweets the tweets of spokespersons on numerous sides of a debate, for example, in order to reflect multiple points of view.

Furthermore, the theory of digital speech acts as exclusionary also advances an understanding of how such vocabulary is the product of both technological affordances and larger societal trends. This thesis has demonstrated how the 'citation needed' phrase, for example, enables a range of digital speech acts that imply rights for readers to have claims on Wikipedia verified. Deploying the phrase, however, requires knowledge of Wikipedia's tagging syntax as well as the appropriate occasions for its deployment according to policy and norms. The 'citation needed' phrase still has power despite the fact that few actually can deploy it because of epistemological changes framing how it is used are more widely recognised. Although there is a large body of work analysing how language has been deployed in order to exclude, there is little analysis of how vocabularies in the digital arena are having a significant affect in the communities in which they are deployed.

Central to Wikipedia's epistemology is the concept of verifiability, a concept originally articulated by the free and open source software movement but one which has become mainstreamed in what has been called the 'age of transparency' (Fung, Graham, & Weil, 2007; Gupta, 2010; Sifry, 2011). This conceptualisation of verifiability articulates the user's ability to look up the claims being made in Wikipedia by following citations to independent sources from which claims were supposedly derived. The assumption behind verifiability as a liberatory act, however, is that all knowledge contained within an article can be verified by independent sources.

Wikipedia has become a reflection of an epistemology which positions the internet as the site of all that is known (or worth knowing), and the process of linking to knowledge in the form of codified information as the primary way in which knowledge can be authoritatively asserted. The framing of the internet as representative of all that is known has force among Wikipedia's readers because it is supported by discourses that are prevalent in the larger socio-technical environment in which readers find themselves. Such discourses, constructed by the Web's dominant actors including Google and the artificial intelligence discourses, tend to equate information with knowledge and result in technologies being developed in order to simulate human knowledge through the processing of information. As Brown and Duguid note,

> People are increasingly eager that their perfectly respectable cache of information be given the cachet of knowledge. Such redefinitions surreptitiously extend the overlapping area where knowledge and information appear as interchangeable terms. (Brown & Duguid, 2000, p. 119)

Wikipedia reinforces this epistemological framing through its assertion that it aims to represent 'the sum of all human knowledge' (Wikimedia Foundation, n.d.), while at the same time imposing limits on what can be represented. A more accurate statement, in keeping with Wikipedia's policies and epistemology, might be that Wikipedia aims to represent the sum of all human *information* since what is represented on Wikipedia is limited to what is already codified. A core feature of Wikipedia's claims to authority is its supposed ability to point to material, *processable* information sources that verify what is represented in the article. Such enactment of verifiability resonates with scientific and academic practices that lend further credibility to cited claims. Thus, Wikipedia, particularly Wikipedia's authority, is both a product of, and produces science, technology and society; that is, Wikipedia is coproduced by science, technology and society.

Furthermore, this thesis demonstrated that power within Wikipedia is held at the edges of the encyclopaedia, so that editors can selectively deploy objects in the form of policy and technological tools in order to advance particular ideologies. According to this framing, the power of central actors including the Wikimedia Foundation, Wikipedia administrators and policy objects, all of which have at least some kind of jurisdiction over the entire network, is limited by the actions of individuals and groups of individuals at the edges of the network. These individuals, both by virtue of their experience or by virtue of their number, are able to influence the representation of knowledge, either because central power brokers do not, or cannot, make editorial decisions or because (in the case of policy objects) rules cannot be formulated for every instance of knowledge work on the site.

This rhizoid distribution of power is not only reflected in peer production projects. Even platforms such as Facebook and Twitter that are corporate owned and subject to terms and conditions determined by the corporations who are responsible for them, exhibit similar power distributions. Like the Wikimedia Foundation, the Twitter corporation has the ability to enforce their ideas about what its users should be able to do and say on the platform, but Twitter's power is limited – both by its users' abilities to deploy Twitter's own tools to revolt against such actions *en masse*, and by the norms and logics of transparency that have become foundational to online discourse that forbids limitations on user speech.

Although the distribution of power on corporate and non-profit platforms is similar, the gap between power exerted at centre and the edges is greater on corporate platforms. The ability of corporations to strictly control the tools by which users access and manage their content on the platform is the greatest source of their power because

corporations can restrict speech to the platform itself. Multiple decentralised tools are available for Wikipedia editors to use in order to filter content on the platform, and significant opportunities are available for editors to build their own tools that connect to the Mediawiki platform.

Facebook and Twitter's APIs (application programming interfaces), on the other hand, are open in the sense that developers can create tools to extract data, for example, but there are more significant limits on what users are able to do to with the data and how they are able to re-publish data on other sites. Despite a more significant power differential, the power for users to represent their views is still effectively held at the edges of the network, and although there are significant discourses that restrict what can be represented, particular sets of users have significantly more power than others, depending on the complexity of the languages used on the sites.

In sum, this thesis contributes to analysis of power in the context of networked knowledge by indicating how power and the identities of the powerful are reconfigured in the context of networked technologies and current discourses around knowledge and expertise. Far from seeing the removal of gatekeepers and the rise of the amateur, this thesis has demonstrated that the power to represent has been reconfigured by dominant actors in the representation of facts about the world.

Power in the context of the decentralised, rhizoid structure of Wikipedia, and indeed, many of today's informational platforms, is the product of the authority of its actors. Authority, both the authority of individuals in the case of powerful Wikipedians, and in concert, in the case of Wikipedia's authority within the larger knowledge environment, is garnered through the enactment of authoritative performances and the uttering of

complex vocabulary and conventions. The authority of Wikipedia and Wikipedians are garnered through the performative acts of citation, through the ability of individual Wikipedians to construct the traces that represent citation, and through the stabilization and destabilization of facts according to the ideological viewpoints of Wikipedia editors.

Theoretically, then, this thesis contributes to current debates about the agency of human actors in the digital environment and the strategies by which power is enacted in the context of the digital. The practical implications of many of the findings also have resonance to debates taking place within the Wikimedia community and help to answer questions that have been raised by journalists investigating Wikipedia's growing power on the internet.

8.2 Implications

There are three important implications of the results of this thesis. The first is that Wikipedia's growing authority (and monopoly) over the representation of digital, structured facts is having an effect on the emergence of other knowledge identities in the wider knowledge system. Of particular importance are the emergence of Wikipedia watchdog groups as those who are disenfranchised by Wikipedia become more visible in mainstream media, and the rise of paid editing (Pinsker, 2015) as a mechanism by which powerful forces in society gain access to Wikipedia's representational tools.

One of the most significant watchdog groups to emerge in recent years is in the website, 'Wikipediocracy'.[59] Founded in March 2012, Wikipediocracy was created by users of the

---

[59] See http://wikipediocracy.com/.

Wikipedia Review, a site that hosted online forums in which ex-Wikipedians (and current Wikipedians) vented about the failings of the encyclopedia. Wikipediocracy bills itself as a site that publishes articles about Wikipedia controversies, conflicts and problems that its authors have identified through varying levels of experience in Wikipedia editing. The authors of Wikipediocracy have consequently become the new revolutionaries against that which they see as Wikipedia's systemic bias and corruption. Wikipediocracy's website frames the group's goals in terms of its role as a defender of the public's right to accurate information.

> We exist to shine the light of scrutiny into the dark crevices of Wikipedia and its related projects; to examine the corruption there, along with its structural flaws; and to inoculate the unsuspecting public against the torrent of misinformation, defamation, and general nonsense that issues forth from one of the world's most frequently visited websites, the "encyclopedia that anyone can edit." (Wikipediocracy, n.d.)

Wikipediocracy articles have, in the past three years of its creation, catalysed reports in the mainstream news media about Wikipedia controversies. Wikipediocracy authors have tipped off journalists about particular stories, and have worked with journalists in order to translate Wikipedia vocabulary and traces for investigative purposes. A Wikipediocracy post that covered Jimmy Wales's relationship to the government of Kazakhstan, for example, was reported in the British Daily Telegraph . Wikipediocracy and the UK Independent, the Telegraph and the Washington Post have run stories in which Wikipediocracy is termed a Wikipedia 'watchdog' (Dewey, 2015b; Duggan, 2014; Gander, 2014). As Wikipedia's authority grows, and more groups feel disenfranchised by its processes, the growth of watchdog groups like Wikipediocracy who act as translators of Wikipedia's complex structures, rules and norms for mainstream media and who begin to give voice to those who feel that they have been excluded from Wikipedia's representational structures will continue.

The need for editors who are able to translate Wikipedia's complex language in order to advance the ideologies of corporate and state interests is also becoming a significant outcome of Wikipedia's growing authority. The Wikimedia Foundation Board developed strict disclosure rules for paid editors of Wikipedia in 2014 ruling, there is evidence to suggest that numerous PR groups continue to edit without disclosing their status as paid editors (Pinsker, 2015). Wikipedia is coming under increasing pressure by groups of paid individuals who have significantly greater resources in order to maintain particularly positive articles for their clients, and the growing complexity of rules, norms and regulations has served to exacerbate this problem.

Without an effective means of regulating paid editing, and the growing complexity of editing practice, it is likely that we will see moves by the WMF towards the deployment of technologies that enable stable identities and digital reputation systems so that there is greater control over who edits Wikipedia in the future. Covert actions are already commonplace on Wikipedia: editors create alternative identities in order to further particular points of view (sockpuppeting) and paid editing without the disclosure of conflicts of interests is possible and therefore regularly uncovered. Covert action by editors is set to continue as Wikipedia's stakeholders attempt to retain the neutrality of the platform in the face of attack by commercial and government forces. Similarly, attempts to make such action transparent through the design of tools such as Wikiscanner[60] will similarly increase which will, in turn, enhance the likelihood of certification procedures for future Wikipedia work.

---

[60] Wikiscanner was a publicly searchable database that linked millions of anonymous edits on Wikipedia to the organizations where those edits apparently originated, by cross-referencing the edits with data on the owners of the associated block of IP addresses (https://en.wikipedia.org/wiki/WikiScanner).

Finally, there are significant implications of the findings relating to the reconfiguration of power in this thesis for the ways in which Wikipedia's diversity problems might be solved. Since power is located at the edges of the encyclopaedia, it is unlikely that policy exceptions can be made for particular types of articles that are underrepresented (such as oral citations) since they would inevitably involve pushback from the Wikipedia community. A more practical solution would be the creation of a separate site for the publishing of oral citations, either a separate organisation or a separate project within the Wikimedia family. This body could fulfil the requirement for sources used on Wikipedia to be published separately from the site, but would be able to develop their own standards for the inclusion of content, particularly with regard to copyright and ethical guidelines for the recording of indigenous knowledge.

Alternatively, the Wikimedia Foundation might provide support for some of the organisations who are already attempting to meet the needs of communities that have remained underserved by the corporate publishing market. Non-profit projects such as Ulwazi,[61] in South Africa already has a platform that enables its members to digitise aspects of local culture but lack the expertise and resources necessary to enable compatibility with Wikipedia. These projects require resources and expertise in order to develop appropriate technical standards for the development of content that can be employed in Wikipedia. Projects like Ulwazi possess the local knowledge necessary to enable active engagement among Wikipedia and local communities regarding copyright and ethical concerns.

---

[61] See http://ulwazi.org/.

It is imperative that Wikipedia finds ways of representing knowledges that remain underserved by the current publishing markets. The implications of a continued reliance on traditional publishing markets will have an impact, not only on Wikipedia's ability to meet its goals, but in the growing number of sites that rely on Wikipedia's data to feed their own databases. Wikipedia is increasingly being recognised as public infrastructure for facts about the world. While it has been suggested that Wikipedia and other peer production projects represent a model that can be used to produce public knowledge products, Wikipedia's systemic biases will need to be countered in order to represent diverse sets of publics' interests.

Wikipedia's decentralised model of development, where individuals and groups decide what they want to work on depending on what is interesting to them, has produced significant inconsistencies, even though Wikipedia as a whole represents more subjects than previous encyclopaedias. Similarly, governments wishing to invest in the peer production model will have to recognise and actively account for Wikipedia's biases when applying the peer production model to other knowledge projects.

8.3 Limitations

This thesis has benefitted from a cosmopolitan approach to fieldwork in which theory is applied to richly textured analysis and a multiplicity of viewpoints. There are, however, three core ways in which the knowledge that is presented here is shaped by the ways in which I decided to follow Wikipedia's multiple networks. My positioning within Wikipedia's networks has undoubtedly affected the ways in which the data has been represented so that this thesis aims at what Haraway calls 'situated knowledge' (1988).

Firstly, although I have accounted for the silence of the subjects of particular facts in the examples of surr's players and the protesters involved in the Egyptian Revolution of 2011, there was a focus in this thesis on those already participating in Wikipedia's representational structures. More depth and scope would be acquired through the interviewing of Wikipedia's silent subjects in future work. Additionally, this thesis did not follow facts as they are employed *after* they have been constructed within Wikipedia. Following the progress of facts that successfully re-enter the source platforms (journalism, academia, social media) from which they were originally derived, would be a useful addition to future work.

Secondly, although mention has been made of the ways in which facts initially represented on Wikipedia move to other platforms such as Wikidata, Google and Facebook, the integrity of facts as they are extracted and (re)presented in different contexts has not been fully interrogated. Following facts as they move between different language versions of Wikipedia (particularly focusing on whether articles are being translated or started from scratch by editors in different language versions) and between from Wikipedia's platform, Wikidata and other platforms would be another fruitful area of research.

Finally, due to the in-depth nature of this study, only two case studies of articles were investigated. The focus of this thesis has been on the interaction between central objects and processes, and the practices and relations between individuals at the article level, and the findings can therefore be generalizable to other articles. Although all articles are subject to the same central processes and policies discussed here, further research into articles that represent different subject areas and knowledge fields may add to the stock of strategies employed by editors in advancing particular perspectives. Source choices

will also be different in articles relating to articles about natural science or medicine, for example. In such cases, further work needs to be done in order to recognise which particular sources are held up as the ideal in the context of those articles, and which are used in practice.

8.4 Future work

In addition to the suggestions for future work indicated in the previous section, there are two further avenues which this research suggests are still particularly important unknowns.

Firstly, the logic of verifiability has been demonstrated to be a central force in shaping of knowledge claims on Wikipedia and there is evidence that other sites are employing similar mechanisms to enhance their authority.[62] Early investigations that I have performed on the use of the 'citation needed' phrase on Twitter indicate new uses for the term beyond Wikipedia. Investigating how discourses embedded in such statements are becoming central features of other media platforms and how they are influencing what are being represented is an important area for research.

Secondly, the rise of new experts and knowledge brokers in the context of automation is a particularly ripe area for research. This thesis introduces the figure of the Wikipedian as a powerful knowledge broker who is heavily aided by technologies that play a role in influencing the scale, anonymity and speed of an individual user's contribution. The

---

[62] Juliette de Maeyer, for example, has investigated the use of hyperlinking in online journalism practice (Maeyer, 2012).

growing importance of Wikidata, however, indicates further reconfigurations of expert identities in the context of an increased reliance on data services in the global economy.

As Wikidata further centralises some of Wikipedia's functionality, new centres of power are emerging that result in unexpected consequences. One of the key areas which could be affected by Wikidata's growing authority is the affordances of the site for discussions only in English where phenomena are represented by only a single item number for all language versions.

Furthermore, understanding changes in the identities of knowledge workers and their roles and responsibilities, as well as how they are interacting with and influencing other knowledge professions such as journalism, science, academia and politics are important areas for future research. This thesis has indicated that there is evidence of Wikipedia possibly encroaching on the work of journalists in the documentation of breaking news events, but further research needs to be done on how the news media is reacting to such changes. Similarly, academics are starting to recognise Wikipedia as an avenue for the dissemination of research findings. Analysing how such communities are engaging with Wikipedia represent some of the most important unasked questions for future research to discover.

8.5   Concluding remarks

The configurations of policies, practices, norms and architecture embedded within Wikipedia's socio-technical system have had a significant impact on the power of individuals to represent their knowledge about the world. Facts that are represented on Wikipedia are becoming part of the infrastructure of the Web so that this impact not only affects what is represented within Wikipedia, but by a myriad of other sites that

make use of Wikipedia data for their own goals. This thesis offers an important answer to the question of how Wikipedia's architecture impacts what knowledge can be represented and how far facts can travel to new users and uses.

Equipped with this understanding, we can begin to recognise how to improve Wikipedia in order to represent the knowledge of underserved groups. Wikipedia will never represent *everyone's* knowledge. The systemic bias embedded within Wikipedia as an encyclopaedia that provides a single definition of a phenomenon, a single representation of the world, means that Wikipedia's bias cannot be entirely countered. Less bias does not mean that we are able to get closer to the truth, but that we are able to get closer to the truths articulated by multiple, diverse communities. Recognising that Wikipedia can never represent all knowledges, but understanding what *can* be done to further its goals for a more consistent coverage of more knowledges, is, however, a project well within reach.

References

AFP. (2011, January 24). Egypt braces for nationwide protests. *Agence France-Presse*. Retrieved from http://www.france24.com/en/20110125-egypt-braces-nationwide-protests/

American University in Cairo. (n.d.). University on the Square: Documenting Egypt's 21st Century Revolution. Retrieved May 14, 2015, from http://digitalcollections.aucegypt.edu/cdm/landingpage/collection/p15795coll7

Anderson, C. W. (2013). *Rebuilding the News: Metropolitan Journalism in the Digital Age*. Philadelphia: Temple University Press.

Arazy, O., Nov, O., & Ortega, F. (2014). The (Wikipedia) world is not flat: On the organizational structure of online production communities. Presented at the Twenty Second European Conference on Information Systems, Tel Aviv, Israel.

Arthur, C. (2013, November 21). Wikipedia sends cease-and-desist letter to PR firm offering paid edits to site. Retrieved July 16, 2015, from http://www.theguardian.com/technology/2013/nov/21/wikipedia-cease-and-desist-pr-firm-offering-paid-edits

Austin, J. L. (1975). *How to Do Things with Words*. (J. O. Urmson & M. Sbisà, Eds.). Oxford: Oxford University Press.

Bach, K. (1998). *Routledge Encyclopedia of Philosophy: Index*. (E. Craig, Ed.). Taylor & Francis.

Baym, N. K. (1999). *Tune In, Log On: Soaps, Fandom, and Online Community*. Thousand Oaks, California: SAGE Publications, Inc.

Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven: Yale University Press.

Benkler, Y., & Nissenbaum, H. (2006). Commons-based Peer Production and Virtue. *Journal of Political Philosophy*, *14*(4), 394–419. http://doi.org/10.1111/j.1467-9760.2006.00235.x

Berry, D. M. (2014). *Critical Theory and the Digital*. New York: Bloomsbury.

Bijker, W. E., Hughes, T. P., & Pinch, T. J. (1987). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge: The MIT Press.

Bloor, D. (1984). The Strengths of the Strong Programme. In *Scientific Rationality: The Sociological Turn* (pp. 75–94). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-94-015-7688-8_3

Brown, J. S., & Duguid, P. (2000). *The Social Life of Information* (1st ed.). Cambridge, MA: Harvard Business Press.

Bruckman, A. (2002). Studying the Amateur Artist: A Perspective on Disguising Data Collected in Human Subjects Research on the Internet. *Ethics and Information Technology*, *4*(3), 217–231.

Bruns, A. (2008). *Blogs, Wikipedia, Second Life, and Beyond: From Production to Produsage*. Oxford: Peter Lang.

Bryant, S. L., Forte, A., & Bruckman, A. (2005). Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work* (pp. 1–10). Sanibel Island, Florida, USA: ACM. http://doi.org/10.1145/1099203.1099205

Burke, P. (2012). *A Social History of Knowledge II: From the Encyclopaedia to Wikipedia*. Cambridge, UK: Polity Press.

Burrell, J. (2009). The Field Site as a Network: A Strategy for Locating Ethnographic Research. *Field Methods*, *21*(2), 181–199. http://doi.org/10.1177/1525822X08329699

Butler, B., Joyce, E., & Pike, J. (2008). Don't look now, but we've created a bureaucracy:

the nature and roles of policies and rules in wikipedia. In *Proceedings of the*

SIGCHI Conference on Human Factors in Computing Systems (pp. 1101–1110).

New York, NY, USA: ACM. http://doi.org/10.1145/1357054.1357227

Cheshire, C., Antin, J., & Churchill, E. (2010). Behaviors, adverse events, and dispositions:

An empirical study of online discretion and information control. *Journal of the*

American Society for Information Science and Technology, *61*(7), 1487–1501.

http://doi.org/10.1002/asi.21346

Clifford, J. (1983). On Ethnographic Authority. *Representations*, (2), 118–146.

http://doi.org/10.2307/2928386

Cohen, J. E. (2006). *Cyberspace As/And Space* (SSRN Scholarly Paper No. ID 898260).

Rochester, NY: Social Science Research Network. Retrieved from

http://papers.ssrn.com/abstract=898260

Collison, R. L. (2015, May 1). Encyclopaedia. In *Encyclopedia Britannica* (online).

Retrieved from http://www.britannica.com/topic/encyclopaedia/Greek-and-

Roman-concepts

Dewey, C. (2015a, February 10). Beyonce fans wrecked Beck's Wikipedia page after the

Grammys. *The Washington Post*. Retrieved from

http://www.washingtonpost.com/news/the-

intersect/wp/2015/02/10/beyonce-fans-wrecked-becks-wikipedia-page-after-

the-grammys/

Dewey, C. (2015b, April 15). The story behind Jar'Edo Wens, the longest-running hoax in

Wikipedia history. *The Washington Post*. Retrieved from

https://www.washingtonpost.com/news/the-intersect/wp/2015/04/15/the-

great-wikipedia-hoax/

Dodge, M., & Kitchin, R. (2005). Code and the Transduction of Space. *Annals of the Association of American Geographers*, *95*(1), 162–180. http://doi.org/10.1111/j.1467-8306.2005.00454.x

Dooley, P. L. (2010). Wikipedia and the Two-faced Professoriate. In *Proceedings of the 6th International Symposium on Wikis and Open Collaboration* (pp. 24:1–24:2). New York, NY, USA: ACM. http://doi.org/10.1145/1832772.1832803

Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge, Mass: The MIT Press.

Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015, January 9). Demographics of Key Social Networking Platforms. Retrieved from http://www.pewinternet.org/2015/01/09/demographics-of-key-social-networking-platforms-2/

Duggan, O. (2014, May 21). Civil servants behind "sickening" Hillsborough slurs identified. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/sport/football/teams/liverpool/10844002/Civil-servants-behind-sickening-Hillsborough-slurs-identified.html

Edgerton, D. (2008). *Shock Of The Old: Technology and Global History since 1900*. London: Profile Books.

Elvebakk, B. (2008). Philosophy Democratized? *First Monday*, *13*(2). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/2091. http://dx.doi.org/10.5210/fm.v13i2.2091

Explain xkcd. (2015, June 29). 285: Wikipedian Protester [blog]. Retrieved from http://www.explainxkcd.com/wiki/index.php/285:_Wikipedian_Protester

Faulkner, P., & Runde, J. (2013). Technological objects, social positions, and the transformational model of social activity. *MIS Quarterly*, *37*(3), 803–818.

Folkenflik, D. (2011, March 3). Clinton Lauds Al Jazeera: 'It's Real News'. *National Public Radio*. Retrieved from http://www.npr.org/sections/thetwo-way/2011/03/03/134243115/clinton-lauds-virtues-of-al-jazeera-its-real-news

Ford, H. (2011). The Missing Wikipedians. In G. Lovink & N. Tkacz (Eds.), *Critical Point of View: A Wikipedia Reader*. Amsterdam: Institute of Network Cultures.

Ford, H., & Geiger, R. S. (2012). "Writing Up Rather Than Writing Down": Becoming Wikipedia Literate. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration* (pp. 16:1–16:4). New York, NY, USA: ACM. http://doi.org/10.1145/2462932.2462954

Ford, H., Sen, S., Musicant, D. R., & Miller, N. (2013). Getting to the Source: Where Does Wikipedia Get Its Information from? In *Proceedings of the 9th International Symposium on Open Collaboration* (pp. 9:1–9:10). New York, NY, USA: ACM. http://doi.org/10.1145/2491055.2491064

Foucault, M. (1980). *Power/knowledge : selected interviews and other writings, 1972-1977*. Hemel Hempstead: Harvester Wheatsheaf Press.

Fuchs, C. (2013). *Social Media: A Critical Introduction*. SAGE Publications.

Fung, A., Graham, M., & Weil, D. (2007). *Full disclosure: the perils and promise of transparency*. Cambridge, UK: Cambridge University Press.

Gander, K. (2014, May 21). Hillsborough Wikipedia posts: Suspected civil servant a Merseyside resident. *The Independent*. Retrieved from http://www.independent.co.uk/news/uk/home-news/hillsborough-wikipedia-posts-suspected-civil-servant-a-merseyside-resident-9413069.html

Geertz, C. (1973). *The Interpretation of Cultures: Selected Essays*. New York: Basic Books.

Geiger, R. S. (2011). The Lives of Bots. In G. Lovink & N. Tkacz (Eds.), *Critical Point of View: A Wikipedia Reader*. Amsterdam: Institute of Network Cultures.

Geiger, R. S., & Ribes, D. (2010). The work of sustaining order in Wikipedia: the banning of a vandal. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 117–126). New York, NY, USA: ACM. http://doi.org/10.1145/1718918.1718941

Geiger, R. S., & Ribes, D. (2011). Trace Ethnography: Following Coordination through Documentary Practices. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on* (pp. 1 –10). http://doi.org/10.1109/HICSS.2011.455

Ghonim, W. (2012). *Revolution 2.0*. New York, US: Houghton Mifflin Harcourt.

Gieryn, T. F. (1983). Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review*, *48*(6), 781–795.

Gieryn, T. F. (2001). Boundaries of Science. In S. Jasanoff, G. E. Markle, J. C. Peterson, & T. Pinch (Eds.), *Handbook of Science and Technology Studies* (pp. 393–443). SAGE Publications.

Gillespie, T. (2010). The politics of "platforms." *New Media & Society*, *12*(3), 347–364. http://doi.org/10.1177/1461444809342738

Gillespie, T. (2014). The Relevance of Algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA; London, England: The MIT Press.

Gillmor, D. (2008). *We the Media: Grassroots Journalism By the People, For the People*. Farnham: O'Reilly Media, Inc.

Glott, R., & Ghosh, R. (2010). *Analysis of Wikipedia Survey Data*. United Nations University MERIT. Retrieved from http://www.wikipediasurvey.org/docs/Wikipedia_Age_Gender_30March%2020 10-FINAL-3.pdf

Goffman, E. (1959). *The presentation of self in everyday life*. London: Penguin Books.

Goodwin, D. (2012, March 19). Bing, Not Google, Favors Wikipedia More Often in Search

    Results [Study]. Retrieved from

    http://searchenginewatch.com/sew/study/2161910/bing-google-favors-

    wikipedia-search-results-study

Graham, M. (2011). Wiki Space: Palimpsests and the Politics of Exclusion. In G. Lovink &

    N. Tkacz (Eds.), *Critical Point of View: A Wikipedia Reader* (pp. 269–282).

    Amsterdam: Institute of Network Cultures.

Graham, M. (2012). *Geography/Internet: Ethereal Alternate Dimensions of Cyberspace or*

    Grounded Augmented Realities? (SSRN Scholarly Paper No. ID 2166874).

    Rochester, NY: Social Science Research Network. Retrieved from

    http://papers.ssrn.com/abstract=2166874

Graham, M., Hogan, B., Straumann, R. K., & Medhat, A. (2014). *Uneven Geographies of*

    User-Generated Information: Patterns of Increasing Informational Poverty (SSRN

    Scholarly Paper No. ID 2382617). Rochester, NY: Social Science Research

    Network. Retrieved from http://papers.ssrn.com/abstract=2382617

Grossman, L. (2006, December 25). You — Yes, You — Are TIME's Person of the Year.

    Time. Retrieved from

    http://content.time.com/time/magazine/article/0,9171,1570810,00.html

Gupta, A. (2010). Transparency in Global Environmental Governance: A Coming of Age?

    Global Environmental Politics, *10*(3), 1–9. http://doi.org/10.1162/GLEP_e_00011

Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the

    Privilege of Partial Perspective. *Feminist Studies*, *14*(3), 575–599.

    http://doi.org/10.2307/3178066

Hauslohner, A. (2011, January 20). After Tunisia: Why Egypt Isn't Ready to Have Its Own

    Revolution. Retrieved from

    http://content.time.com/time/world/article/0,8599,2043497,00.html

Hern, A. (2015, January 23). Wikipedia votes to ban some editors from gender-related

    articles. Retrieved June 11, 2015, from

    http://www.theguardian.com/technology/2015/jan/23/wikipedia-bans-

    editors-from-gender-related-articles-amid-gamergate-controversy

Hersch. (2012, March 25). Verifiability vs. truth [Blog]. Retrieved from

    http://wikipediocracy.com/2012/03/25/verifiability-vs-truth/

Hess, D. (2001). Ethnography and the Development of Science and Technology Studies.

    In P. Atkinson, A. Coffey, S. Delamont, J. Lofland, & L. Lofland (Eds.), *Handbook of*

    Ethnography (pp. 234–245). London: SAGE Publications

Hjorth, I. (2013). *Networked cultural production: Filmmaking in the Wreckamovie*

    community (Doctoral dissertation) University of Oxford, UK.

Hill, B. M., & Shaw, A. (2013). The Wikipedia Gender Gap Revisited: Characterizing

    Survey Response Bias with Propensity Score Estimation. *PLoS ONE*, *8*(6),

    e65782. http://doi.org/10.1371/journal.pone.0065782

Hine, C. (2008). Virtual Ethnography: Modes, Varieties, Affordances. In *The SAGE*

    Handbook of Online Research Methods (pp. 257–270). London, England: SAGE

    Publications, Ltd. Retrieved from http://srmo.sagepub.com/view/the-sage-

    handbook-of-online-research-methods/n14.xml

Isin, E. F., & Ruppert, E. (2015). *Being digital citizens*. London, UK ; New York: Rowman &

    Littlefield.

Jackson, S. J., Edwards, P. N., Bowker, G. C., & Knobel, C. P. (2007). Understanding

    infrastructure: History, heuristics and cyberinfrastructure policy. *First Monday*,

    12(6). Retrieved from

    http://firstmonday.org/ojs/index.php/fm/article/view/1904

Jasanoff, S. (Ed.). (2004). *States of knowledge: the co-production of science and social*

    order. London; New York: Routledge.

Jemielniak, D. (2014). *Common Knowledge?: An Ethnography of Wikipedia*. Stanford, California: Stanford University Press.

Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide*. New York: New York University Press.

Johnson, N. (2008, May 15). Wikipedia Traffic Grows 8,000% in 5 Years Due to Search Referrals. Retrieved from http://searchenginewatch.com/sew/study/2054211/wikipedia-traffic-grows-years-due-search-referrals

Jovchelovitch, S. (2006). *Knowledge in Context: Representations, Community and Culture*. New York, NY: Routledge.

Jurkowitz, M. (2011, February 5). Events in Egypt Trigger Record Coverage. Retrieved from http://www.journalism.org/index_report/pej_news_coverage_index_january_31_february_6_2011

Keegan, B. C. (2013). A History of Newswork on Wikipedia. In *Proceedings of the 9th International Symposium on Open Collaboration* (pp. 7:1–7:10). New York, NY, USA: ACM. http://doi.org/10.1145/2491055.2491062

Keen, A. (2008). *The Cult of the Amateur: How blogs, MySpace, YouTube and the rest of today's user-generated media are killing our culture and economy* (New Ed edition). London: Nicholas Brealey Publishing Ltd.

Kildall, S., & Stern, N. (2011). Wikipedia Art: Citation as Performative Act. In G. Lovink & N. Tkacz (Eds.), *Critical Point of View: A Wikipedia Reader* (pp165-190). Amsterdam: Institute of Network Cultures.

Kildall, S., & Stern, N. (n.d.). Wikipedia Art | a collaborative performance and public intervention [Digital artwork archive]. Retrieved July 21, 2015, from http://wikipediaart.org

Kitchin, R., & Dodge, M. (2011). *Code/space: software and everyday life*. Cambridge, Mass:
    MIT Press.

Kollock, P. (1999). The Economies of Online Cooperation: Gifts and Public Goods in
    Cyberspace. In *Communities in Cyberspace* (pp. 239, 220). New York: Routledge.
    Retrieved from
    http://www.sscnet.ucla.edu/soc/faculty/kollock/papers/economies.htm

Kreiss, D., Finn, M., & Turner, F. "The Limits of Peer Production: Some Reminders from
    Max Weber for the Network Society." *New Media & Society.* 13(2) 243-259,
    2011. http://doi:10.1177/1461444810370951

Kriplean, T., Beschastnikh, I., McDonald, D. W., & Golder, S. A. (2007). Community,
    Consensus, Coercion, Control: Cs*W or How Policy Mediates Mass Participation.
    In *Proceedings of the 2007 International ACM Conference on Supporting Group
    Work* (pp. 167–176). New York, NY: ACM.
    http://doi.org/10.1145/1316624.1316648

Kumar, K. (2013, May 23). Official language or national language? *The Hindu*. New Delhi.
    Retrieved from http://www.thehindu.com/features/friday-review/official-
    language-or-national-language/article4743140.ece

Lam, S. (Tony) K., Uduwage, A., Dong, Z., Sen, S., Musicant, D. R., Terveen, L., & Riedl, J.
    (2011). WP:clubhouse?: An exploration of Wikipedia's gender imbalance (p. 1-
    10). ACM Press. http://doi.org/10.1145/2038558.2038560

Larsen, J. (2008). Practices and Flows of Digital Photography: An Ethnographic
    Framework. *Mobilities*, *3*(1), 141–160.
    http://doi.org/10.1080/17450100701797398

Latour, B. (1987). *Science in Action: How to Follow Scientists and Engineers Through
    Society*. Boston, MA.: Harvard University Press.

Latour, B. (1991). The Berlin Key or How to Do Things with Words. In P. Graves-Brown (Ed.), *Matter, Materiality and Modern Culture* (pp. 10–21). London; New York: Routledge.

Latour, B. (1992). Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts. In W. Bijker & J. Law (Eds.), *Shaping Technology/Building Society: Studies in Sociotechnical Change* (pp. 225–258). Cambridge, Mass.: The MIT Press. Retrieved from http://www.bruno-latour.fr/articles/article/050.html

Latour, B. (1993). *We have never been modern*. Boston, MA.:Harvard University Press.

Latour, B., & Woolgar, S. (1979). *Laboratory Life: The Construction of Scientific Facts*. (J. Salk, Ed.). Beverly Hills, CA: SAGE Publications.

Lam, S., K., Uduwage, A., Dong, Z., Sen, S., Musicant, D. R., Terveen, L., & Riedl, J. (2011). WP:clubhouse?: An exploration of Wikipedia's gender imbalance (p. 1-10). ACM Press.

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation* (1ST edition). Cambridge, England; New York, USA: Cambridge University Press.

Leadbeater, C., & Miller, P. (2004). *The Pro-Am Revolution: How Enthusiasts are Changing Our Society and Economy*. London: Demos.

Lessig, L. (2009). *Remix: making art and commerce thrive in the hybrid economy*. New York: Penguin Books.

Lewis, S. C. (2012). The Tension Between Professional Control and Open Participation. *Information, Communication & Society*, *15*(6), 836–866. http://doi.org/10.1080/1369118X.2012.674150

Leyne, J. (2011, January 17). No sign Egypt will take the Tunisian road. *BBC News*. Cairo. Retrieved from http://www.bbc.co.uk/news/world-middle-east-12202937

Lievrouw, L. A. (2014). Materiality and Media in Communication and Technology Studies: An Unfinished Project. In T. L. Gillespie, P. J. Boczkowski, & K. A. Foot

(Eds.), *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: The MIT Press.

Loveland, J., & Reagle, J. (2013). Wikipedia and encyclopedic production. *New Media & Society*. http://doi.org/10.1177/1461444812470428

Lucassen, T., & Schraagen, J. M. (2010). Trust in Wikipedia: how users trust information from an unknown source. In *Proceedings of the 4th workshop on Information credibility* (pp. 19–26). New York, NY, USA: ACM. http://doi.org/10.1145/1772938.1772944

Lucassen, T., & Schraagen, J. M. (2013). The influence of source cues and topic familiarity on credibility evaluation. *Computers in Human Behavior*, *29*(4), 1387–1392. http://doi.org/10.1016/j.chb.2013.01.036

Luyt, B. (2011). The nature of historical representation on Wikipedia: Dominant or alterative historiography? *Journal of the American Society for Information Science and Technology*, *62*(6), 1058–1065. http://doi.org/10.1002/asi.21531

Luyt, B. (2012). The inclusivity of Wikipedia and the drawing of expert boundaries: An examination of talk pages and reference lists. *Journal of the American Society for Information Science and Technology*, *63*(9), 1868–1878. http://doi.org/10.1002/asi.22671

Luyt, B., & Tan, D. (2010). Improving Wikipedia's credibility: References and citations in a sample of history articles. *Journal of the American Society for Information Science*, *61*(4), 715–722. http://doi.org/10.1002/asi

Maasen, S., & Weingart, P. (2006). *Democratization of Expertise?: Exploring Novel Forms of Scientific Advice in Political Decision-Making*. Dordrecht, The Netherlands: Springer Science & Business Media.

Mackenzie, A. (2006). *Cutting Code: Software and Sociality*. New York, USA: Peter Lang Publishers.

MacKenzie, D. A., & Wajcman, J. (Eds.). (1985). *The social shaping of technology* (2nd ed). Buckingham [Eng.]. Philadelphia: Open University Press.

Maeyer, J. D. (2012). The Journalistic Hyperlink. *Journalism Practice*, *6*(5-6), 692–701. http://doi.org/10.1080/17512786.2012.667273

Marcus, G. E. (1995). Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography. *Annual Review of Anthropology*, *24*(1), 95–117. http://doi.org/10.1146/annurev.an.24.100195.000523

Markham, A., & Buchanan, E. (2012). *Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0)*.

Marshall, G. (Ed.). (1994a, June 1). ethnography. *The Concise Oxford Dictionary of Sociology* (p. 158). Oxford; New York: Oxford Paperbacks.

Marshall, G. (Ed.). (1994b, June 1). participant observation. *The Concise Oxford Dictionary of Sociology* (p. 380). Oxford; New York: Oxford Paperbacks.

Mathis-Lilley, B. (2015, March 13). Edits of Wikipedia Pages for Eric Garner, Stop-and-Frisk Were Made From NYPD Headquarters. *Slate*. Retrieved from http://www.slate.com/blogs/the_slatest/2015/03/13/nypd_wikipedia_edits_eric_garner_sean_bell_amadou_diallo.html

Messer-Kruse, T. (2012, February 12). The "Undue Weight" of Truth on Wikipedia. *The Chronicle of Higher Education*. Retrieved from http://chronicle.com/article/The-Undue-Weight-of-Truth-on/130704/

Miller, J. (2014, July 25). Wikipedia blocks "disruptive" page edits from US Congress. *BBC News*. Retrieved from http://www.bbc.co.uk/news/technology-28481876

Morgan, M. S. (2010a). *How well do facts travel? Final project report*. Retrieved from http://www.lse.ac.uk/economicHistory/Research/facts/Home.aspx

Morgan, M. S. (2010b). Travelling Facts. In *How Well Do Facts Travel?: The Dissemination of Reliable Knowledge* (pp. 3–39). Cambridge, UK: Cambridge University Press.

Newton, S. I. (1959). *The correspondence of Isaac Newton*. Published for the Royal

Society at the University Press. London, England.

Niederer, S., & van Dijck, J. (2010). Wisdom of the crowd or technicity of content?

Wikipedia as a sociotechnical system. *New Media & Society*, *12*(8), 1368–1387.

http://doi.org/10.1177/1461444810365297

Nielsen, F. A. (2007). Scientific citations in Wikipedia. *First Monday*, *12*(8). Retrieved

from http://firstmonday.org/ojs/index.php/fm/article/view/1997

Oreskes, N., & Conway, E. M. (2011). *Merchants of Doubt: How a Handful of Scientists*

Obscured the Truth on Issues from Tobacco Smoke to Global Warming (Reprint

edition). New York, NY: Bloomsbury Press.

Pentzold, C., & Seidenglanz, S. (2006). Foucault@ Wiki: first steps towards a conceptual

framework for the analysis of Wiki discourses. In *Proceedings of the 2006*

international symposium on Wikis (pp. 59–68). ACM New York, NY, USA.

Pinsker, J. (2015, August 11). The Covert World of People Trying to Edit Wikipedia—for

Pay. *The Atlantic*. Retrieved from

http://www.theatlantic.com/business/archive/2015/08/wikipedia-editors-for-

pay/393926/

Ramesh, R. (2015, April 21). Grant Shapps accused of editing Wikipedia pages of Tory

rivals. *The Guardian*. Retrieved from

http://www.theguardian.com/politics/2015/apr/21/grant-shapps-accused-of-

editing-wikipedia-pages-of-tory-rivals

Ramos, D. (2015, June 1). Wikipedia editors honor Caitlyn Jenner's identity.

BostonGlobe.com. Retrieved from

https://www.bostonglobe.com/opinion/2015/06/01/wikipedia-editors-honor-

caitlyn-jenner-identity/kMBuLrQJYZUZefLBVeVd7H/story.html

Reagle, J. (2010). *Good Faith Collaboration: The Culture of Wikipedia*. Cambridge, MA: The MIT Press.

Reagle, J. (2013). "Free as in sexist?" Free culture and the gender gap. *First Monday*, *18*(1). http://doi.org/10.5210/fm.v18i1.4291

Reich, E. S. (2011). Online reputations: Best face forward. *Nature News*, *473*(7346), 138–139. http://doi.org/10.1038/473138a

Rosenzweig, R. (2006). Can History Be Open Source? Wikipedia and the Future of the Past. *The Journal of American History*, *93*(1), 117–146. http://doi.org/10.2307/4486062

Ryle, G. (1968). *The thinking of thoughts*. Saskatoon, Canada: University of Saskatchewan.

Sandvig, C. (2013). The Internet as Infrastructure. In *The Oxford Handbook of Internet Studies*. Oxford, UK: Oxford University Press.

Schiff, S. (2006, July 31). Know It All. *The New Yorker*, (July 2006). Retrieved from http://www.newyorker.com/magazine/2006/07/31/know-it-all

Schneider, J., Samp, K., Passant, A., & Decker, S. (2013). Arguments About Deletion: How Experience Improves the Acceptability of Arguments in Ad-hoc Online Task Groups. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (pp. 1069–1080). New York, NY, USA: ACM.

Sen, S. W., Ford, H., Musicant, D. R., Graham, M., Keyes, O. S., & Hecht, B. (2015, April). Barriers to the localness of volunteered geographic information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2015*.

Shapin, S. (1984). Pump and Circumstance: Robert Boyle's Literary Technology. *Social Studies of Science*, *14*(4), 481–520. http://doi.org/10.1177/030631284014004001

Shirky, C. (2009). *Here comes everybody : how change happens when people come together*. London, UK: Penguin.

Sifry, M. L. (2011). *Wikileaks and the Age of Transparency*. OR Books.

Silverwood-Cope, S. (2012, February 8). Wikipedia: Page one of Google UK for 99% of searches. Retrieved from http://www.intelligentpositioning.com/blog/2012/02/wikipedia-page-one-of-google-uk-for-99-of-searches/

Singhal, A. (2012, May 16). Introducing the Knowledge Graph: things, not strings [blog]. Retrieved from http://googleblog.blogspot.com/2012/05/introducing-knowledge-graph-things-not.html

Society for the Diffusion of Useful Knowledge. (1833). *The Penny Cyclopaedia of the Society for the Diffusion of Useful Knowledge*. London, UK: Knight Publishers.

Star, S. L. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, *43*(3), 377–391. http://doi.org/10.1177/00027649921955326

Stelter, B. (2011, January 31). Al Jazeera English Finds an Audience. *The New York Times*. Retrieved from http://www.nytimes.com/2011/02/01/world/middleeast/01jazeera.html

Stephens, M. (2013). Gender and the GeoWeb: Divisions in the production of user-generated cartographic information. *GeoJournal*, *78*(6), 981–996. http://doi.org/10.1007/s10708-013-9492-z

Sundin, O. (2011). Janitors of knowledge: constructing knowledge in the everyday life of Wikipedia editors. *Journal of Documentation*, *67*(5), 840–862. http://doi.org/10.1108/00220411111164709

Tkacz, N. (2012). Wikipedia and the Politics of Openness (Doctoral dissertation) The University of Melbourne, Australia.

Teplitskiy, M., Lu, G., & Duede, E. (2015). Amplifying the Impact of Open Access:

Wikipedia and the Diffusion of Science. *arXiv:1506.07608 [physics]*. Retrieved

from http://arxiv.org/abs/1506.07608

Thomas Macaulay. The Miscellaneous Writings and Speeches of Lord Macaulay., Volume

4 (of 4) by Thomas Babington Macaulay (1831). House of Commons, Britain.

Retrieved from http://www.gutenberg.org/files/25903/25903-

h/files/2170/2170-h/2170-h.htm#2H_4_0006

Trist, E. L., & Bamforth, K. W. (1951). Some Social and Psychological Consequences of the

Longwall Method of Coal-Getting An Examination of the Psychological Situation

and Defences of a Work Group in Relation to the Social Structure and

Technological Content of the Work System. *Human Relations*, *4*(1), 3–38.

http://doi.org/10.1177/001872675100400101

TTFNROB. (2015, March 5). Who are the Zooniverse community? We asked them...

[Weblog]. Retrieved from http://blog.zooniverse.org/2015/03/05/who-are-

the-zooniverse-community-we-asked-them/

van Dijck, J. (2013). *The Culture of Connectivity: A Critical History of Social Media*. Oxford,

UK: Oxford University Press.

Viegas, F., Wattenberg, M., Kriss, J., & van Ham, F. (2007). Talk Before You Type:

Coordination in Wikipedia. In *Hawaii International Conference on System

Sciences, Proceedings of the 41st Annual* (Vol. 40, p. 1298). IEEE.

Weinberger, D. (2011). *Too Big to Know: Rethinking Knowledge Now that the Facts Aren't

the Facts, Experts are Everywhere, and the Smartest Person in the Room is the

Room*. London, UK: Basic Books.

Weller, K., Dornstädter, R., Freimanis, R., Klein, R. N., & Perez, M. (2010). Social software

in academia: Three studies on users' acceptance of Web 2.0 services. In

Proceedings of the 2nd Web Science Conference (WebSci10), Retrieved May 29,

2013 from: http://www.phil-

fak.uniduesseldorf.de/fileadmin/Redaktion/Institute/Informationswissenschaft

/weller/websci 10_submission_62.pdf Wennerås (p. 341343).

Wells, H. G. (1938). *World Brain* (Reprint edition). Freeport, N.Y: Ayer Co Pub.

Wikimedia Foundation. (2009, May 21). Wikimedia community approves license

migration. Retrieved from http://blog.wikimedia.org/2009/05/21/wikimedia-

community-approves-license-migration/

Wikimedia Foundation. (2011). *Wikipedia Editors Study: Results from the Editor Survey,*

April 2011. Wikimedia Foundation. Retrieved from

https://upload.wikimedia.org/wikipedia/commons/7/76/Editor_Survey_Repor

t_-_April_2011.pdf

Wikimedia Foundation. (n.d.). Vision statement. Retrieved July 24, 2015, from

https://wikimediafoundation.org/wiki/Vision

Zuckerman, E. (2010, March 24). Makmende's so huge, he can't fit in Wikipedia.

Retrieved from

http://www.ethanzuckerman.com/blog/2010/03/24/makmendes-so-huge-he-

cant-fit-in-wikipedia/

Primary sources

'Wikipedia:Autopatrolled'. English Wikipedia autopatrolling process. Retrieved from https://en.wikipedia.org/w/index.php?title=Wikipedia:Autopatrolled&oldid=657274011

'Common knowledge'. (2015, July 23). Retrieved from https://en.wikipedia.org/w/index.php?title=Common_knowledge&oldid=672762505

Encyclopedia. (2015, August 12). Retrieved from https://en.wikipedia.org/w/index.php?title=Encyclopedia&oldid=675716237

'Wikipedia:Be Bold'. English Wikipedia editing guideline. Retrieved from https://en.wikipedia.org/w/index.php?title=Wikipedia:Be_bold&oldid=666074629

'Wikipedia:Core content policies'. English Wikipedia information page. Retrieved from https://en.wikipedia.org/w/index.php?title=Wikipedia:Core_content_policies&oldid=671517499

'Wikipedia:HotCat'. English Wikipedia project page. Retrieved from https://en.wikipedia.org/w/index.php?title=Wikipedia:HotCat&oldid=673977553

'Wikipedia:Identifying and using primary and secondary sources. English Wikipedia essay. Retrieved from https://en.wikipedia.org/w/index.php?title=Wikipedia:Identifying_and_using_primary_and_secondary_sources&oldid=669692807

'Wikipedia:Identifying reliable sources'. English Wikipedia content guideline. Retrieved from https://en.wikipedia.org/w/index.php?title=Wikipedia:Identifying_reliable_sources&oldid=673989715

'Wikipedia:List_of_policies'. Summary of English Wikipedia policies. https://en.wikipedia.org/w/index.php?title=Wikipedia:List_of_policies&oldid=669228761

'Wikipedia:No original research'. (WP:NOR). English Wikipedia policy. Retrieved from https://en.wikipedia.org/w/index.php?title=Wikipedia:No_original_research&oldid=672653340

Wikipedia:Reliable Sources/Noticeboard. English Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Wikipedia:Reliable_sources/Noticeboard

'Wikipedia:Truth, not verifiability'. English Wikipedia essay. Retrieved from https://en.wikipedia.org/w/index.php?title=Wikipedia:Truth,_not_verifiability&oldid=603392200

User:Going Batty. [Going Batty] user page. Retrieved from https://en.wikipedia.org/wiki/User:GoingBatty

'Wikipedia:Verifiability'. English Wikipedia policy. Retrieved from
https://en.wikipedia.org/w/index.php?title=Wikipedia:Verifiability&oldid=672654354

'Wikipedia:Verifiability, not truth'. English Wikipedia essay. Retrieved from
https://en.wikipedia.org/w/index.php?title=Wikipedia:Verifiability,_not_truth&oldid=672421172

'Wikipedia:What Wikipedia is Not'. English Wikipedia policy. Retrieved from
https://en.wikipedia.org/w/index.php?title=Wikipedia:What_Wikipedia_is_not&oldid=674256264

'Wikipedia:Wikipedia Manual of Style'. English Wikipedia guideline. Retrieved from
https://en.wikipedia.org/w/index.php?title=Wikipedia:Manual_of_Style&oldid=674001305

Appendix A: Participant information sheet

<u>The practices of knowledge construction on Wikipedia</u>

Researcher
Heather Ford is a DPhil (PhD) student at the Oxford Internet Institute, University of Oxford, OX1 3AJ, United Kingdom. Contact information: <u>heather.ford@oii.ox.ac.uk</u>

Purpose of this study
This project aims to explore how Wikipedians construct knowledge in the process of daily practice and routines.

Why are you being asked to participate?
As someone who is an important stakeholder in one of the Wikipedia case studies under analysis, your experience and thoughts are very valuable for this study.

What will participation in this study involve?
I would like to have conversations with you about your involvement with Wikipedia, what work you do, who you work with, and what motivates you. Basically, I would like to hear your stories. These conversations/interviews can be conducted via Skype, telephone, emails, online chat or in person and will take approximately 30 minutes to 45 minutes. Recordings will be collected during the interview in order to record the detail of what was discussed. This data will be stored securely, on a password-protected drive. If you withdraw from this study, the files containing your data will be destroyed.

What are your rights?
You can change your mind at any time. If you no longer wish to be part of this study, you simply send an email informing me of that decision. No questions asked.

What if there is a problem?
If you have a concern about any aspect of this project, please contact me and I will do my very best to answer your query. If you remain unhappy and wish to make a formal complaint, please contact the Research Ethics Committee at the University of Oxford (ethics@socsci.ox.ac.uk).

Are there any risks or benefits?
I do not believe that there are any risks involved in participating in this study. You may benefit personally in some way by articulating and reflecting on your experiences and contributing to knowledge that will be used to improve Wikimedia's development around the world.

Questions?
Please let me know if you have further questions related to this study, or want to know more before you make your decision.

Appendix B: Consent form

<u>Consent form</u>

I confirm that I have read this form and received a copy of it. I have had all my questions answered to my satisfaction and understand:

- Data collected as part of this study will remain confidential;

- Participants may choose whether their comments will be attributed to their 'real name', their Wikipedia pseudonym or anonymised using pseudonyms developed specifically for the study;

- Participation is voluntary and one can withdraw from the study at any time. In this case all data will be destroyed;

- Audio recordings will be made during the interview process (if audio interviews occur, otherwise email interviews will be stored). This data will be stored securely;

- All data collected about, or identifying, participants will be stored securely on a password protected drive.

I agree to take part in this study.

Printed name: _____

Signature: _____

Date: _____

Consent can be given by:

a.   Signing a physical copy of this form and returning it to the researcher, or

b.   Signing an electronic copy of this form. Fill out the form and send to the researcher via email. Please mention in the accompanying email that you have read the consent form and personally signed it.