Climate Change Coverage in the English and Nordic Wikipedia

Anda Bordieanu

Supervisor: Florian Maximilain Meier

[137,777 characters with spaces, 57.4 pages]

Table of contents

Climate Change Coverage in the English and Nordic Wikipedia

Anda Bordieanu[1]

[1] Aalborg University Copenhagen

Information Studies Master

A. C. Meyers Vænge 15, 2450 Copenhagen, Denmark

abordi20@student.aau.dk

Abstract. There is a wide consensus that climate change represents the challenge of the modern world. Its effects are becoming more and more visible and efforts for mitigating and adapting to these changes are of high priority. Public awareness is an important factor when considering mitigation policies and one platform that people use to inform themselves about climate change is Wikipedia. The largest online multilingual encyclopedia reached a permanent spot in the public eye, who use it to attain knowledge. Therefore, this study investigates the climate change coverage in the English and Nordic Wikipedia by conducting semi-structured interviews with Wikipedia's editors and collecting data via Wikipedia's API. The results presented showcased that there are information asymmetries between English and Nordic Wikipedia, where substantial discrepancies in the content can be noticed. The English Wikipedia is more thoroughly maintained compared to the Nordic editions, with addition of higher quality content in regards to the topic of climate change.

Keywords: climate change, Wikipedia and climate change, Wikipedia and climate change representation, Wikipedia and knowledge gaps, Wikipedia and multilingualism

1 Introduction

A lot of attention is circled around global warming as it represents a serious environmental challenge that our society has to face (Nordhaus, 2007). The terms *global warming*, *carbon emissions, climate change* and *sustainability* have become the buzzwords of the decade (Scoones, 2007). There is a wide consensus that climate change is a complex and uncertain problem that can threaten the life of different species and can bring major changes to society (Moser, 2010; Tol, 2009). It is a challenge that our and future generations will have to face, learn how to live with it, and also how to mitigate it (Brügger et al., 2015).

According to the Intergovernmental Panel on Climate Change (IPCC), the last three decades have shown warmer temperatures compared to any other decades since 1850 (IPCC, 2007, p.40). There is overwhelming evidence that the earth's climate is currently undergoing a change (Blennow & Persson, 2009). Global temperatures are expected to rise by 4-celsius degrees by the year 2100 and the climate conditions are going to be altered (Thuiller, 2007). Climate change is one of the most pressing topics in the contemporary world and it is described as the top global risk in the following 10 years by the World Economic Forum (Varini et al., 2020). The climate change effects on the environment are gradually more visible. Different phenomena such as fires, droughts and floods are now stronger and they happen frequently (Rolnick et al., 2022). Efforts for mitigating and adapting to the new climate conditions are recommended and are crucial, thus the need to tackle these effects of climate change is seen as a priority on an international level (Budžytė & Balžekienė, 2018). However, addressing climate change without public involvement will not be possible (Luís et al., 2018).

In 2019, around 1.9 million protesters joined the worldwide youth strike for climate change. The global climate strike was established by Greta Thunberg, the young climate activist to raise attention to this topic and encourage actions for mitigations (Boulianne et al., 2020). Moreover, even celebrities are starting to raise awareness of the climate crisis (Doyle et al., 2017). In 2016, the world-famous Hollywood actor, Leonardo DiCaprio, accepted his Oscar by giving a speech about climate change. This had an impact on social media engagement on this subject, as the tweets directly related to this were 636% higher than expected on the day DiCaprio advocated for climate change (Leas et al., 2016). There was also a substantial increase in Google searches, more specifically 261% more searches right after DiCaprio held his speech, and this topic maintained such significant interest even four days after the speech occurred. These searches marked the third-highest point ever registered for climate change or global warming queries on Google (Leas et al., 2016). Thus, it is highlighted that celebrities

understand the immediate need for public awareness and engagement with this topic (Park, 2019).

Climate change is seen as an environmental crisis, thus it is a topic that should be accessible to a wide audience. The way in which climate change is communicated to the public can have an impact on their perception of this issue. Moreover, the coverage in the media on climate change can shape behavioural choices, such as reducing the carbon footprint or even participating in climate activism (Whitmarsh & Capstick, 2018). As the internet is highly used by millions of people for different reasons, it also has great potential for scientific communication (Lörcher & Taddicken, 2017). People can easily search for information regarding the topics that are of interest to them, and one platform that is used for this purpose is Wikipedia (Ford & Wajcman, 2017).

1.2 Problem field

There is a general understanding that the main cause of climate change is human activities which led to an increase in greenhouse gas emissions (Blennow & Persson, 2009; Henderson et al., 2015). However, many are still denying climate change and even though they have knowledge about it, they often reinterpret that information thinking that it happens due to natural causes, for example, the rising in the solar system activities (Esteves & Cukierman, 2012; Esteves Gonçalves da Costa, Bernardo & Cukierman, 2019; Häkkinen & Akrami, 2014; Luís et al., 2018). The public denial of climate change and the causes behind it represents a challenge to create and implement efficient policies for mitigation (Krange et al., 2019). With the fast and high availability of information on the internet, it is now easy to manipulate public perception with false news or claims (Diggelmann et al., 2020).

Wikipedia, the largest online encyclopedia on the web is hosted by the Wikimedia Foundation and encompasses a multitude of articles about different topics. The main characteristic of the encyclopedia is that anyone that has access to the internet can become an editor (Keegan et al., 2012;Viégas et al., 2007). Wikipedia is one of the most popular websites nowadays as it receives around 14 billion page views per month (Dang & Ignat, 2020; Hinnosaar et al., 2021). It has the capability to provide facts regarding different scientific subjects. (Baker et al., 2020, p.8; Kincaid et al., 2021). People around the world demand access to climate change knowledge and they search on Wikipedia to find reliable information, thus the accuracy and reliability of its data are important. Therefore, the online encyclopedia is a powerful platform for enabling sharing of verified and accurate facts about this subject (Baker et al., 2020, p.8; Kincaid et al., 2021).

However, the collaborative aspect of the online encyclopedia is a concern when it comes to the quality and trustworthiness of its articles (Bassani & Viviani, 2019; Sepehri-Rad & Barbosa, 2015). Controversies about different subjects have in fact occurred on Wikipedia, and back in 2006, the editors of an article could not decide on the introduction of a photo in an article (Forte et al., 2009). A more recent article in the encyclopedia that has gotten a lot of attention from the public is concerned with describing the current situation that is happening in Ukraine. The article is called *2022 Russian invasion of Ukraine* and it contains the events which are currently occurring while the war is ongoing. On the 24th of February, the first day of the Russian invasion, the article was edited each minute by different users (Wikipedia, 2022). The growth of Wikipedia in the number of editors has become a challenge when it comes to its governance. It was shown that this growth is leading to an increase in the cost of coordination activities within the encyclopedia (Forte & Bruckman, 2008). This can raise questions regarding subjects such as climate change and whether they are or are not affected by the controversy and bias that can be found on Wikipedia (Forte et al., 2009). To make sure that the content of Wikipedia is accurate, over the years, within its own group of contributors, WikiProjects was formed (Bassani & Viviani, 2019). These communities aim to evaluate the articles and eliminate biases (Bassani & Viviani, 2019; Forte et al., 2009).

In 2012, Google introduced Knowledge Graph, which made it possible to display knowledge from third parties sources on the Google Search Engine Results Page. When a user searches on Google, the SERP includes a Knowledge Panel, which extracts general information on the topic usually from Wikipedia (Rothshild et al., 2019). Wikipedia's popularity rose and its pages take up the top search engine results, which is the source of most of Wikipedia's traffic (Menchen-Trevino & Hargittai, 2011). Wikipedia is not only the largest encyclopedia available online, but it is also the largest multilingual encyclopedia. However, the non-English content on Wikipedia is just a mere fraction of the English one (Roy et al., 2020). Topics that may be covered in the English Wikipedia, might be represented differently in the other languages due to the bias or preferences of the editors. The open and free content of Wikipedia makes it the main source of information when users search for subjects in their native language, yet there exist information asymmetries between English Wikipedia and other language editions (Filatova, 2009; Roy et al., 2022). For instance, the number of articles in Hindy represents only 2% when compared to the English edition, even though Hindi is the second spoken language in the world (Roy et al., 2022).

Because Wikipedia can be edited by anyone who has access to the internet, it is different from a paper-based source of information. An assumption formed around Wikipedia is that it is based upon the concept of the *wisdom of the crowds*, where the collective opinion of wider independent masses is deemed superior to that of a single expert (Santana & Wood, 2009). This idea that the content on Wikipedia is improved by having more editors was only partially supported in a study. The research highlighted that quality will improve only if good coordination techniques are used by the contributors. It pointed out that even an open-source project like Wikipedia needs coordination, when too many editors are involved, peer-to-peer communication is unsuccessful (Kittur & Kraut, 2008). With this in mind, the purpose of this research is to investigate the global issue of climate change and its portrayal in Wikipedia by looking into the WikiProject Climate Change. The project is currently available in the English edition, but taking into account that people are used to searching in their local language, we will focus on the Swedish, Danish and Norwegian languages (Nordic countries) to understand if the information presented in the English Wikipedia can also be found in the Nordic countries. This leads to the following problem statement.

Climate change is a topic that gained a lot of attention in the last decade especially since it was proved that human-induced climate change does exist. The effects of the sudden change in the weather patterns have caused a lot of controversies both in the media and in the online landscape. As Wikipedia is one of the main sources of information online, the goal of this project is to investigate how climate change is represented in the largest online encyclopedia, with a focus on the English and the Nordic language editions of it. Investigating these languages could highlight information asymmetry between English and Nordic Wikipedia.

To investigate the problem, the following research questions will be answered.

RQ1: How can Wikipedia make sure that climate change subject is covered at the same level on the English and Nordic Wikipedia?

RQ2: How many articles about climate change are in the English Wikipedia compared to the Nordic language editions and what is their quality?

RQ3: How often do climate change articles get visited and what sparks more interest in the readers?

To answer these research questions, a literature review will be conducted as the first stage of this research. This will offer a deeper understanding of Wikipedia as a source of information and its previous topic controversies or biases found. Moreover, the term climate change and its representation on Wikipedia will be better grasped. During the second stage, the methodology for data collection and analysis will be selected. The

first two research questions will depend upon the Wikipedia data which will be extracted by connecting to their API, whereas the final question will be answered by utilising the interview results. For data analysis, Rstudio will be used along with thematic analysis for understanding the data.

2 Literature search and review

This section highlights the literature related to the problem field. First, the literature search process is described, followed by the literature regarding climate change, Wikipedia and its background, and continuing with describing Wikipedia and the climate change project, finally ending with the asymmetries found on Wikipedia across different languages.

2.1 Literature Search

This study aims to investigate how climate change topic is covered on Wikipedia, focusing on the English and Nordic countries' editions. Researchers must be able to build and relate their research to previous studies, to account for the focus of the investigation, and also to determine the basis of the study (Snyder, 2019).

To conduct a literature review, four distinct types can be followed, *traditional or narrative literature review, systematic literature review, meta-analysis and meta-synthesis* (Cronin et al., 2008).

This research makes use of the concepts presented in the traditional literature review to understand and conduct the synthesis of different works and scholarly articles related to the aim of the investigation. The main role of the traditional literature review is to provide a foundation for understanding the current information and knowledge already studied in the interested area of research (Cronin et al., 2008). All the literature presented in this chapter was selected by searching on different online databases, such as JSTOR, Google Scholar, AAU library, Scopus, and IEEE, but also on an online library, that contains multiple books both academic and non-academic, called Perlego.

Before starting the search process, first, the keywords or phrases that are related to the concepts of the study were identified by using a strategy called *building blocks*. This strategy can be used only after the purpose of the research was identified as it requires dividing the topic of the study into different themes (Boren & Moxley, 2015). By using this technique the keywords and synonyms describing the subject of research were identified and it helped in finding even more literature that could be summarised, for instance, Wikipedia and Climate Change, Information asymmetry and Wikipedia

language editions, Wikipedia and Climate Change coverage, showcased in figure 1. This strategy can be a great help to recognise all the different keywords that can return more articles to be included in the synthesis.

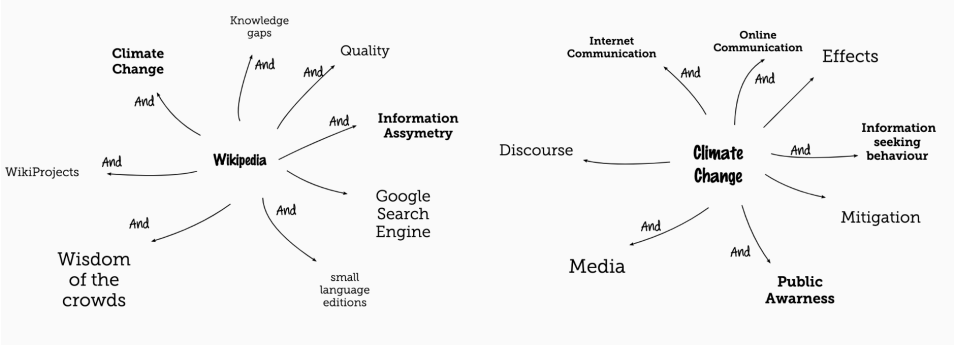


Figure 1. Building Blocks diagram

Aside from the building blocks strategy, another technique utilised to search for related studies was the pearl growing method or snowball technique. Following this strategy, once the main articles for this research were identified, their bibliographies and citations were checked to find more literature for understanding the topic of this paper better (Wohlin, 2014). The list of the initial main articles found, their results and reasons for including them in this study can be seen in Appendix A.

The snowball and building blocks techniques provided a considerable amount of articles to be synthesized. To establish the important and useful information presented that should be mentioned in the research the *PQRS (Preview, Question, Read, Summarise)* method was used (Cronin et al., 2008). First, via the preview stage, the abstracts of the research papers were analysed as it provides a comprehensive summary into understanding if the articles are related to the main focus of the research. Second, the purpose of the studies, the methods used, and the results were questioned, to end with summarizing the key findings of the selected literature. Thus, the process of literature review consisted of three stages, deciding on the search strategy, keywords used, and conducting the search. These stages were iterated until no literature could be found. The articles included in this section are the ones which were found to be relevant to the research field. First, the importance of climate change mitigation and adaptation in relation to the public perception and their information-seeking behaviour was synthesised. Second, Wikipedia and its background were reviewed to understand its structure and culture. Third, the portrayal of climate change subject in Wikipedia was

scrutinised, and lastly, the information asymmetries that were found in Wikipedia were described.

2.2 Climate Change and Information Seeking Behaviour

Climate change is by no doubt a global issue. The IPCC recommended using our efforts for finding different ways to adapt to and mitigate the notable effects of climate change (Blennow & Persson, 2009). Some of the most important measures are carbon sequestration and the use of renewable energy, such as wind, solar and geothermal energy sources (VijayaVenkataRaman et al., 2012).

Numerous countries started to take initiative, and most of the agreements between countries call for limiting the global average temperature increase below 2 degrees, but also for striving to limit it below 1.5 degrees. This stabilisation requires that the rate of increase of greenhouse gas emissions will be equal to the rate at which natural resources can erase them (Ussiri & Lal, 2017, p. 293).

With the Paris agreement, 197 countries have signed to use their efforts to limit the increase of C02 and stay within a carbon budget of 400- 800 GtCO2 (Fankhauser et al., 2022). This means that the emissions will reach the highest point in 2030, but by 2050 they will fall to net zero. Alongside countries, more and more companies are committing themselves to reach net-zero by mid-century (Fankhauser et al., 2022). In April 2020, Shell was one of the latest oil-and-gas leaders to declare a form of a commitment to achieve net-zero (Diringer & Perciasepe, 2020).

Implementing different mitigation initiatives is challenging as it also requires involvement from the public (Diringer & Perciasepe, 2020). If consumers show interest in this topic, then by adopting a more sustainable way of living, they can substantially reduce greenhouse gas emissions from domestic consumption and transport (Semenza et al., 2008). Thus, climate change policies depend on public support due to the changes that individuals have to implement in their lifestyles (Lee et al., 2015). A study found that awareness regarding climate change is important. People who are aware that human-induced climate change is happening, are more likely to support mitigation changes rather than adaptive policies or no action taken at all (Eichhorn et al., 2020, p.21). Moreover, people who relate climate change to human activities are more than four times as likely to favour different mitigation policies in comparison with those who believe that climate change is happening due to natural causes. Additionally, there is also a correlation between mitigation action and impact awareness. Those who forecast the negative global effects of climate change are also the ones in favour of supporting

mitigation action (Eichhorn et al., 2020, p.37). However, climate change is a difficult topic to communicate.

The public assimilates knowledge about different scientific topics, including climate change from a multitude of channels, such as media with television, news, and documentaries representing a common source of information (Anderson, 2011). However, in the last years, the channels utilised for climate change communication have expanded, as it is attempted to reach wider audiences via more channels (Schäfer, 2012). In fact, a new environment for climate change communication is formed by the Internet (Koteyko et al., 2015). Thus, the internet is considered the most suitable platform to communicate climate change information (Kažys, 2018).

Web Search Engines are one of the most important sources of information retrieval for any web-based system (Jamali & Asadi, 2010). Back in 2012, it was shown that for any day, half of the adults who were using the Internet, were using a search engine (Purcell et al., 2012). If in the past, there was no dominant search engine, nowadays, Google is described as the most popular search engine, due to its precision in retrieving relevant search results (Jamali & Asadi, 2010; Waller, 2011). Moreover, search engine market share in Europe showed that Google holds 91.66% of the market and in another work, it was highlighted that nine out of ten minutes that a person spends on the internet is used on a Google website, either the search engine of Google, Google Books, or Gmail (Waller, 2011). A study that investigated the seeking behaviour of people concerning topics such as climate change, highlighted that people search for more information on Google regarding climate change because of their personal experiences. Google searches increased when extreme heat in the summer was occurring, or there were long periods without rain. Therefore, weather anomalies which people are experiencing are strongly connected with their information-seeking behaviour (Lang, 2014).

A study conducted in 2021 focused on millennials and how this generation seeks information about climate change. Millennials are people born between 1982 and 1999, and their upbringing was highly influenced by the internet (Luqman, 2021). 70% of the adults between 18 and 34 years expressed concerns about global warming, whereas only 56% of adults aged 55 and onward showed the same worry. This generation is seen as the most diverse group of adults concerned with various global issues, with climate change being the biggest worry for them, followed by wars, conflicts and inequality (Luqman, 2021; Ross et al., 2019). The study showcased that millennials do not use standard sources of information, such as TV, radio or paper-based media when searching for climate change information. The internet is the main source of information for this generation (Luqman, 2021).

As a response to the global environmental crisis, climate change education is seen as an important constituent as it can raise awareness and enforce climate-conscious values to the young generations. The quality of climate change education is inter-connected with the sources of information and their quality used by the teachers to educate future generations (Puttick & Talks, 2021). Numerous studies that investigated the sources of information used by teachers from 2010 and onward, found that the Internet is the dominant source (Higde et al., 2017; Iordanou & Constantinou, 2014; Johnson, 2011; McNeal et al., 2017). Moreover, a study developed a learning environment which guided teachers to specific sources of information, one of them being the online encyclopedia, Wikipedia (Iordanou & Constantinou, 2014).

Wikipedia's web pages are appearing among the top results on search engines and previous studies highlighted that people tend to read the first articles showcased in the SERP (Kazys, 2018). Thus, a study investigated if there is a link between the web search frequency and the page views of Wikipedia. The findings of the study showed a correlation of 0.92 and a concordance rate of 0.54. It highlighted that Wikipedia's page views can be a way to reproduce web search trends (Yoshida et al., 2015).

According to Eichhorn et al. (2020) knowledge about climate change and its effects on the globe, as well as its effects on each individual is an important factor when it comes to mitigation action (Eichhorn et al., 2020, p.46). As it was shown, the internet nowadays represents the main source of information retrieval, as it provides an easy way for different people to find out more about their topics of interest, for example, climate change. Wikipedia is one of the most important sources of information and its articles are among the top results on the most popular search engine utilised, therefore, the following sections focus on the online encyclopedia.

2.3 Wikipedia and its Background

Wikipedia is an unparalleled example of mass collaboration (Wilkinson & Huberman, 2007). Since it was first launched in 2001 by Jimmy Wales, Wikipedia is now the largest online encyclopedia and the sixth most visited website worldwide. It receives over 85 million visitors every month in the US alone (Vetter, 2018). The encyclopedia records each month a total of 15 billion page views, and traffic coming from more than 1.5 billion unique devices from around the globe (Redi et al., 2020). People go to Wikipedia for different trivial reasons, either to check on their favourite celebrities' personal lives or to gain knowledge about more serious facts (Warncke-Wang et al., 2015).

Wikipedia is by far the most successful collaborative project that aims to gather all the knowledge and share it freely with every person in the world (Vetter, 2018). Most people know Wikipedia as a collection of articles, but the encyclopedia has an

infrastructure made of different web pages. For instance, each article has a category, which makes it easy to search for different topics as well as to discover missing topics (Johansson & Lindberg, 2019).

The editors, the people that transformed Wikipedia as we know it today, collaborate and form a culture having different norms and chains of behaviour. There can be a real person behind a user account, or not, and the users can choose if they want to disclose their identity or stay anonymous. Currently, there are 77 million users registered on Wikipedia, but this number does not reflect the active users (Johansson & Lindberg, 2019). In fact, according to Morgan et al. (2013), Wikipedia needs more diversified editors. A diverse community of contributors will increase the quality and completeness of the online encyclopedia (Morgan et al., 2013; Schneider et al., 2014). Curiously, 25% of the active users found on Wikipedia are bots, however, the handful of human active editors are responsible for the edits related to the articles which are most read (Priedhorsky et al., 2007). Each language edition has its own community of contributors and they are self-governed. Even if Wikipedia can be edited by anyone, exceptions do apply. Certain users can have more power, they are the admins and they can decide to ban a user or block them from editing a specific article (Johansson & Lindberg, 2019).

Previous studies highlighted that there exist interdependencies between Wikipedia and Google Search (McMahon et al., 2017; Vincent et al., 2019). Many query searches generate more Wikipedia links as results than any other website on the Google SERP. This showcases that the online encyclopedia plays an important role in the ways through which Google Search achieves its aim to address the information needs of the public (Vincent & Hecht, 2021). Interdependencies exist also in regards to news and Wikipedia. It was highlighted that the most edited and viewed articles in the encyclopedia are the ones in regards to current events happening in the world and that readers are influenced by the media coverage to read Wikipedia (Keegan et al., 2012; (Singer et al.,2017). Furthermore, Wikipedia covers even natural disasters as a study showed that when the Tōhoku catastrophe occurred, a lot of editing was conducted in the encyclopedia in the time surrounding the event (Keegan et al., 2011). Additionally, Reddit, an online network of communities, also plays a role in Wikipedia's page traffic. The users' posts related to the articles in the encyclopedia, were shown to increase the page views on Wikipedia (Moyer et al., 2015). As such, Wikipedia attained a permanent place in the public eye, which uses it to gain knowledge (Vetter, 2018).

One feature that distinguishes Wikipedia from any other encyclopedias is the fact that everyone can edit it. However, a challenge that Wikipedia faces is attaining and retaining new editors that can quickly understand its norms. Failing to provide a good onboarding for its new editors can lead to low-quality articles, and frustration among the new editors (Narayan et al., 2017). Over the years Wikipedia has received criticism

regarding systematic bias, containing falsehoods, being edited with inappropriate purposes or just for the sake of disruption (Dunn et al., 2018). Moreover, the openness that comes with Wikipedia is often related to misinformation and vandalism (Das et al., 2022). Back in 2005, an article containing false information was missed by Wikipedia's New Page Patrol, a group of editors who checks the quality of newly created articles. An anonymous user edited the article about John Seigenthaler, Sr., a distinguished American journalist and tied him to the assassination of Kennedy. The mistake was only caught after five months when it was brought to the attention of Jimmy Wales by the journalist himself (Ayers et al., 2008, p.51). Other recent studies showcased that even though vandalism is still occurring on Wikipedia, it is very fast corrected and when it comes to controversial articles, they can be locked and protected from editing when other editors consider it necessary (Tran, 2009). Moreover, all of the online encyclopedia's editors should follow a set of collaboratively created guidelines, and policies when they write articles on Wikipedia (Redi et al., 2019). These are *Neutral Point of View,* where editors should represent all the views about a topic, by making sure they exclude their own biases. *Verifiability,* where other users that read the articles can check the reliability of the sources that are reflected in the text. *No original research* relates to the fact that no article should contain personal research (Saez-Trumper, 2019). Moreover, it was shown that Wikipedia is one of the most comprehensive sources of information which has a vast coverage regarding topics in nearly all fields of knowledge (Mesgari et al., 2015). The accuracy of Wikipedia was compared to the one of Encyclopedia Britannica, and no substantial difference in the accuracy of both was found (Priedhorsky et al., 2007).

2.4 Wikipedia and Climate Change

The way in which scientific knowledge about climate change is represented through genres in the public sphere has an impact in regards to understanding the topic but also when it comes to policies for mitigating it. One of the sources utilised for information seeking about scientific topics in the contemporary world is Wikipedia (Cooke, 2021).

The information about climate change is mostly represented in the English language, but people are used to search for their topic of interest in their native language (Kažys, 2016). Additionally, they tend to focus more on the first search top results when searching for information on the Web, as the higher positions are favourable (Vincent et al., 2019). Wikipedia appeared among the top three positions for search results about climate change in multiple countries, but in the nordic countries, such as Norway, Denmark, Finland and Sweden, Wikipedia appeared in the first position (Kažys, 2018). Moreover, according to Vincent et al., (2019), 80% of the top three results are represented by Wikipedia's articles. Google is depending on the encyclopedia more than it depends on any other website to reach its goals and offer access to information to all users. Thus the online encyclopedia is an important source of knowledge about climate change, and its contents should be kept up to date so that it remains relevant (Kažys, 2016).

Writing on Wikipedia is done in a collaboratively way and editors can communicate via talk pages. In fact, each article written in the encyclopedia has a talk page where comments regarding any modifications or disagreements can be aligned (Poudat et al., 2016). These pages often contain information regarding how editors decide upon Wikipedia's policies' representation in the specific article. Sometimes, editors have different views on the topic, and these can lead to conflicts (Cooke, 2021). According to Poudat et al., (2016) global warming and climate change are some of the topics that did generate dispute on the French Wikipedia. This happened in the English Wikipedia as well. The Global Warming article was under discussion as it was said that it does not represent the views of the climate change sceptics. The editor backed its idea by stating that the purpose of his article is not to share opinions on global warming but to represent peer-reviewed research (Cooke, 2021). Thus, in this case, it can be seen how Wikipedians made use of the neutral point of view and verifiability policy to make sure that the information shared is accurate (Baker et al., 2020, p.16).

As Wikipedia can be edited by anyone on the internet, no matter if that person is an expert in the climate change field or has any academic background, a study conducted on the Portuguese language edition sought to investigate if there is any room for climate sceptics (Esteves Gonçalves da Costa, Bernardo & Cukierman, 2019).

Climate sceptics are often supported by institutions linked to the oil industry or other institutions opposed to any regulations that could affect their business, thus denying anthropogenic climate change. The research showcased that even though vandalism has occurred in the article about global warming, those statements were quickly corrected. Controversies did happen, and discussions took place between editors as some had different opinions. However, Wikipedia's policies were used to explain why the statements which were related to climate scepticism were left out, thus settling the controversy. One editor pointed out that the encyclopedia should prioritize mainstream science. The study concluded that in that article there was no room for scepticism (Esteves Gonçalves da Costa, Bernardo & Cukierman, 2019).

Apart from editors using talk pages to edit a certain article, if editors would like to work with others on a specific topic, they can form a WikiProject. Usually, these projects aim to maintain the quality of Wikipedia's articles. There were some WikiProjects related to climate change, but in 2019 the WikiProject Climate Change was launched. These projects can have multiple subgroups or task forces (Baker et al., 2020, p.18). The WikiProject Climate Change has currently four task forces, Africa task force, climate justice task force, agriculture task force and biodiversity task force. This project is a descendant of WikiProject Environment and has 72 active members (Wikipedia). In general, WikiProjects and their task forces rate articles based on different criteria. Their quality and importance are taken into account and the highest grade they can have is *FA or feature article.* This refers to articles that were substantially checked and reviewed impartially and covered on a comprehensive note the topic. The lowest rating an article can have is *stub*, which means that it superficially describes the subject (Dang & Ignat, 2020). The other grades that can be assigned to articles are seen in figure.2. Regarding importance, articles are rated as *top, high, medium or low* (Dunn et al., 2018).

TABLE 2 WikiProject article grading scheme[28]

Class	Criteria
FA	**Feature article** **Reader's experience**: Professional, outstanding and thorough; a definitive source for encyclopedic information. **Editing suggestions**: No further content additions necessary unless new information becomes available.
FL	**Feature list** **Reader's experience**: Professional standard; it comprehensively covers the defined scope, usually providing a complete set of items, and has annotations that provide useful and appropriate information about those items. **Editing suggestions**: No further content additions necessary unless new information becomes available.
A	**A-class** **Reader's experience**: Very useful to readers. A fairly complete treatment of the subject. A non-expert in the subject would typically find nothing wanting. **Editing suggestions**: Expert knowledge may be needed to tweak the article, and style problems may need solving.
GA	**Good article** **Reader's experience:** Useful to nearly all readers, with no obvious problems; approaching (but not equalling) the quality of a professional encyclopedia. **Editing suggestions**: Some editing by subject and style experts is helpful; comparison with an existing featured article on a similar topic may highlight areas where content is weak or missing.
B	**B-class** **Reader's experience**: Readers are not left wanting, although the content may not be complete enough to satisfy a serious student or researcher. **Editing suggestions**: A few aspects of content and style need to be addressed. Expert knowledge may be needed.
C	**C-class** **Reader's experience**: Useful to a casual reader, but would not provide a complete picture for even a moderately detailed study. **Editing suggestions**: Considerable editing is needed to close gaps in content and solve cleanup problems.
Start	**Start article** **Reader's experience**: Provides some meaningful content, but most readers will need more. **Editing suggestions**: Providing references to reliable sources should come first; the article also needs substantial improvement in content and organisation.
Stub	**Stub article** **Reader's experience**: Provides very little meaningful content; may be little more than a dictionary definition. **Editing suggestions**: Any editing or additional material can be helpful. The provision of meaningful content should be a priority.

Figure 2. Article Grading Scheme on the English Wikipedia. Taken from Baker et al., (2020)

Back in July 2020, the WikiProject Climate Change had a total of 2340 articles, out of which 5 feature articles, and 22 articles were deemed as having top importance. The rest of the articles were start (917), stub (454), or C class (588). Moreover, the top most edited articles were *2019-20 Australian Bushfires Season*, *Greta Thunberg*, and *List of active coal fired power stations in Turkey* (Baker et al., 2020, p.22). These can be seen in figure 3.

FIGURE 4 Climate change articles by quality and importance, as of July 2020[31]

Quality	Importance						
	Top	High	Mid	Low	NA	???	Total
FA	1	1	2	1			5
GA		13	13	6			32
B	13	43	48	40		4	148
C	22	124	124	215	2	101	588
Start	7	88	154	315	10	343	917
Stub		8	42	150	11	243	454
List	1	7	8	5		6	27
NA	3		1		160		164
Assessed	47	284	392	732	183	697	2,335
Unassessed				1		4	5
Total	47	284	392	733	183	701	2,340

FIGURE 5 WikiProject Climate Change's 'hot articles', on 6 February 2020[32]

152 edits	2019–20 Australian bushfire season
100 edits	Greta Thunberg
48 edits	List of active coal fired power stations in Turkey
39 edits	Bushfires in Australia
36 edits	Elizabeth Wathuti
28 edits	Sustainable energy
23 edits	Air conditioning
21 edits	Climate Museum
18 edits	Xiye Bastida
17 edits	Eren Holding

Figure 3. Climate Change articles their quality and importance and the most edited articles in 2020 taken from Baker et al., (2020)

Silva (2021), a climate change disinformation specialist from BBC, ran an article on climate denial and conspiracy issues in the small language editions of Wikipedia and found that in the Swahili, Kazakh and Belarusian language editions, it is highlighted that scientists are not agreeing on climate change causes (Silva, 2021). Moreover, on the Chinese Wikipedia, it appears that one of the reasons that the global temperature has increased is due to solar activities, whereas on the Croatian edition, most of the content is promoting conspiracy theories while also denying the science that supports climate change (Silva, 2021). Wikimedia Foundation is relying on its editors to make sure that the content is accurate, as they do not get involved directly when it comes to volunteers respecting its editorial policies (McDowell & Vetter, 2020; Silva, 2021).

Additionally, in a podcast called *Wikimedia and Climate Change Misinformation: It could happen here*, which had Alex Stinson, the Wikimedia Foundation senior programme strategist, as a guest, the climate change topic and its coverage on Wikipedia was discussed (Evans, 2021). He mentioned that the neglect of content is one of the few things that stand behind the misinformation regarding climate change across all the language editions. Moreover, as climate change is a subject that

requires a deep understanding of scientific communication and one needs to know where to look for information, people tend to follow a dominant idea in the narrative coming from the media. Thus, manipulation of the discussion surrounding climate change, where the media is inclined to follow a certain view can also influence the editor's opinion (Evans, 2021). According to Cook (2022), even a small piece of misinformation regarding climate change is successful in lowering people's belief and acceptance of climate change, which also has an impact on supporting the policies regarding climate change misinformation (Cook, 2022).

The neglect of content on this topic can delay the action needed by the public because they do not have the right information. Wikipedia is trying to portray all the perspectives of climate change, as long as a reputable source defends a statement. However, sometimes these sources, even though they seem to be well respected, are founded by the fossil fuels industry, and these do not represent the truth of climate change (Evans, 2021). According to Dunlap & McCright (2011), the fossil fuel industry realised from its early days that recognizing the fact that climate change is real and the main cause is the carbon emissions, could be a threat to their economical growth. As such, they launched climate scepticism to a new level, where they undermined the entire field of science, and induced doubt and uncertainty about climate change and its causes (Dunlap & McCright, 2011; p.146, 144-160).

As Wikipedia is an open-source project, it is prone to disinformation. The language editions, which have a vast editorial community, are able to quickly fight against misinformation, as editors can quickly catch the false statements (Evans, 2021). Moreover, the larger language editions have bots developed to patrol the content and reverse any doubtful piece of information. Regarding the small editions, these usually have a small community of editors, which are not active all the time, thus, they are vulnerable to vandalism. In more languages, there is a need for the climate change representation to be relevant to their countries, as different regions need different information, especially when it comes to the adaption to climate change (Evans, 2021).

2.5 Wikipedia and Information Asymmetry

Wikipedia aims to offer free access to all the knowledge available to the public, this means that everyone, no matter where they come from should be able to read and contribute to the encyclopedia. Thus, to make sure that everyone can read its content, Wikipedia covers numerous languages, as it has more than 300 active languages (Vrandecic, 2018). Even though the encyclopedia is multilingual, there exist content gaps across different Wikipedia editions (Miquel-Ribé & Laniado, 2020). The knowledge that is represented in other language editions, is a mere fraction when put in contrast with the English edition. A study highlighted that the English edition of

Wikipedia contains 13.8% of all articles on the encyclopedia, with more than 4.9 million articles. However, the Chinese Wikipedia has a total of 827,273, and the Arabic edition has 373,064 articles (Kim et al., 2016). Moreover, the Hindy language, the third most spoken in the world, comprises 147,000 articles (Roy et al., 2022). Only a few language editions have more than one million articles, 59 language editions have over 100,000 articles, and 137 have more than 10,000. (Vrandecic, 2018; Wulczyn et al., 2016).

The uneven distribution between the number of articles between the English Wikipedia and the other languages is not the only asymmetry present. There is also a significant gap among the editors, which is one of the reasons that there is in the first place such disparity between the language editions (Roy et al., 2022). For instance, the English edition has 6.4 times more active editors than the German Wikipedia, which is one of the most active editions. Approximately 38.4% of all edits on Wikipedia were made in the English language (Kim et al., 2016). This can have consequences on the coverage of different topics across all languages. The writing of articles in different editions in the online encyclopedia is done in a spontaneous manner, and it was shown that cultural and geographical factors have an impact on the distribution of content. Most of the time, the topic of interest in a specific country is only represented in the local language edition (Roy et al., 2020). This is the case not only for the small editions but also for the English edition, even though it is the largest. For example, the English Wikipedia only covers 51% of articles that exist in the German language (Kim et al., 2016). Furthermore, even the similarity between the same topics presented in more language editions can vary. Some articles can be simply translated, others are written from scratch by new editors specific to the language they represent (Barrón-Cedeno et al., 2014)

The asymmetry regarding coverage is also present in Scandinavian countries. According to Bick (2014), the Swedish Wikipedia has 6 times more content than the Danish edition. Although the contrast between the two editions in terms of content can be due to the creation of articles coming from bots, the Swedish edition has 3.2 more edits, highlighting that it has more human editors than the Danish edition (Bick, 2014). In 2020, the Swedish Wikipedia was the second-largest edition, with 3.7 million articles, followed by the German language (2.4 million articles), and the French and Dutch editions with 2.2 million and respectively 2 million articles (Baker et al., 2020, p.13).

Many topics found in Wikipedia are covered in more languages, however, their content can vary widely. Some articles could be written independently, while others could be translated from other languages. This leads to different language editions representing different aspects of a subject (Barrón-Cedeno et al., 2014). Wikipedia was also described as a corpus, as it is the largest encyclopedia covering multiple languages.

A study extracted two different types of corpora from Wikipedia, one called not-aligned, and the other aligned. Not-Aligned corpora were defined as those articles that can be found in two languages on the same topic but the content does not correspond between the two, while aligned-corpora have the same coverage in both languages. The findings showed that the size of the aligned corpora was smaller than the non-aligned one (Otero et al., 2011).

To overcome this challenge impose by asymmetry and information inequality, Google founded a contest to motivate students to contribute to the Swahili edition, while in Germany the government supported the creation of Wikipedia articles by allocating funds (Kim et al., 2016). Moreover, the Wikimedia Foundation started a program in 2010 to motivate students to contribute to Wikipedia, called Wiki Education Project (WikiEd). Teachers are highly encouraged to incorporate in their classes, writing assignments related to Wikipedia. This represents one of the largest interventions which encompasses 63,000 students that contributed to 83,000 articles (Li et al., 2020). Currently, Wikimedia Foundation's Research team is working to identify and establish a knowledge gap index to assist the decision-making process across Wikimedia. They have identified three main categories of knowledge gaps, such as content, readers and contributors (Redi et al., 2020). Retaining and integrating new contributors is necessary for Wikipedia, as well as having them contribute to more language editions (Li et al., 2020; Schneider et al., 2014). This can lead to the climate change topic being relevant and more accessible. All language editions must have climate change and its issues covered, as well as locally important events regarding adaption and mitigation (Baker et al., 2020, p.7).

To find out how climate change is represented in the English Wikipedia but also in the Nordic countries edition, this research will strive to discover any asymmetries between the editions regarding the content, by investigating the WikiProject Climate Change. The literature reviews provided key insights for this study. Wikipedia is an important communication channel for climate change due to its dependencies with Google SERP. It was proven to be affected by a wide range of knowledge gaps, such as contributor, readership and content gaps, and information asymmetries were discovered between different language editions. Moreover, Wikipedia struggles in attaining and retaining new editors, which are vital for its growth and accuracy.

3 Research Design and Methods

This section describes the research design, including the possible approaches in regards to data collection and analysis, the philosophical assumption regarding this study, followed by the methods used for collecting data regarding the climate change representation on Wikipedia. Moreover, the analysis of the data collected using both qualitative and quantitative methods is explained.

3.1 Possible Research Approaches

The final direction of this research was determined when the strategy for the study was selected. However, before settling upon the research design for this paper, other strategies were considered. Some quantitative and qualitative methods were investigated such as surveys, case studies, and experiments.

Experiments are used to examine the relationship between an independent and dependent variable. The focus is put on the probability of a change occurring in an independent variable, and how this change affects the dependent variable (Saunders et al., p.190). In this research, we could have made use of experiments and tested how climate change articles (dependent variable) on Wikipedia were affected by a mediatised event related to it occurring, for instance, the IPCC report release. Using experiments would have meant changing the research questions into hypothesises, as we would be interested in testing the different relationships between the variables of interest.

Apart from experiments, the survey was another method of interest. This strategy which makes use of questionnaires for data collection is a popular choice in research as they offer the possibility of collecting standardised data from a considerable amount of people (Saunders et al.,2019, p.193). Questionnaires are easy to administer as they can be sent online via email, social media or other communication channels. Moreover, it is a convenient tool for participants as they can answer the questions when their time permits (Bryman, 2012, p.670). However, the problem of no response should be taken into account. Most surveys will have several non-responses as people can refuse to participate in the research. It was highlighted that there can be noticed a decline in the response rates of surveys regarding social research (Bryman, 2012, p.199). Thus, the survey strategy with questionnaires as a method for data collection could have been used instead of semi-structured interviews, however, the data collected would not have been as extensive.

When it comes to qualitative research strategy, grounded theory was scrutinised to understand if it could be used. With grounded theory, the focus is on the data collection and analysis techniques. This strategy aims to establish theoretical clarifications of social relations in a variety of contexts, such as management and business (Saunders et al., 2019, p.208). Grounded theory is usually applied by the researcher collecting data from an interview or observation and instead of continuing collecting data, first, the analysis of the initial interview will be conducted. It is recommended to analyse the data as close as possible to the time of collection when conducting this type of research. This will highlight emerging new cases that could be considered in the research. Thus, the participants in the research are selected based on previous analysis (Saunders et al., 2019, p.208). This was not selected in this research as it is a rigorous approach which requires a lot of time and consideration.

Regarding the qualitative research method, focus groups were highly considered as a method for data collection. In this case, the focus would have been on how the attendees interact with each other and how they respond to different opinions regarding the climate change issues on Wikipedia (Bryman, 2012, p.501). However, this method was not selected as it would not have been feasible. Wikipedia's editors are hard to get in touch with and it would have been equally harder to have them all present in a focus group at the same time.

All these considerations regarding different strategies and method techniques were studied before settling on mixed-method research, which is better presented in the next sub-chapters. Collecting quantitative data from Wikipedia regarding climate change coverage in the Nordic and English editions, and completing the results with the findings from the qualitative semi-structured interviews was seen as the best option. This could provide a better view of the problem by also understanding the opinions of those who are currently editing on the WikiProject Climate Change.

3.2 Research Philosophy

To investigate climate change coverage on Wikipedia by comparing different language editions imposes certain beliefs and assumptions about generating new knowledge. In every stage of research, there are assumptions about the reality encountered also called ontology, about human knowledge, epistemology, but also about the values that researchers share that can influence the way a study goes, called axiology. Moreover, the researchers can also have different standpoints, either objectivists or subjectivists. Objectivists believe the truth about reality can be grasped via measurable facts, and they are trying to remain as neutral to the research as possible, whereas a subjective researcher understands that there can be different opinions and narratives that should be studied in detail to comprehend different social realities. Having an understanding of

these set of presumption shape the way research is investigated and puts the basis of a well-thought research philosophy that can highlight the methods to be used for data collection and analysis (Saunders et al., 2019, p. 196). When epistemology, ontology, axiology and methodology are put together they are forming a paradigm. Therefore, a research paradigm can be defined as a lens used by the researcher to conceptualise the study and inspect the methodological perspective of the research and its methods (Kivunja & Kuyini, 2017).

There are numerous paradigms, each one having its own set of assumptions regarding ontology, epistemology, and methodology. Commonly, each paradigm comes with one research method, either quantitative or qualitative. However, one paradigm, *pragmatism*, involves using more than one type of research method when collecting the data (Saunders et al., 2019, p. 196). This paradigm is characterised firstly by the fact that knowledge in research is gathered by utilising methods that are deemed suitable by the researcher. Secondly, from the ontology, there is not a single reality, a single truth, as people have their own understanding of reality. Thirdly, the belief that stands behind this paradigm is that knowledge can be acquired by utilising mixed methods when conducting the research (Kivunja & Kuyini, 2017; Rahi, 2017). Mixed methods research represents a study in which both quantitative and qualitative methods are used (Bryman, 2012, p. 628). This research employs the pragmatism paradigm to fully get an understanding of climate change representation in Wikipedia. Therefore the methods chosen for this paper are the ones that are seen as the most appropriate, and the reality of the climate change coverage in the encyclopedia can be explained from two perspectives: One being the data that can be extracted from Wikipedia in which an objective view is portrayed, and the second one, the qualitative data that can introduce a new reality to the study, by following a subjective approach.

3.3 Research Design

The main part of a research design is to fully comprehend the methods used for data collection and analysis, as well as to define the time horizon for the research (Saunders et al., 2019, p. 172).

This research will focus on the number of articles found on the WikiProject, their quality, and their views which can highlight what sparks more interest in the readers and if there is information asymmetry between the Nordic editions and the English Wikipedia. Therefore, the research follows a cross-sectional study as the data regarding all the above-mentioned variables were collected at a single point in time. Usually, a cross-sectional study entails as a research strategy, survey research. However, qualitative

methods can also be used as they often employ a form of a cross-sectional study, for instance, semi-structured interviews (Bryman, 2012, p. 59).

The mixed methods approach was selected for this study as it allows the researcher to tackle complex research questions, find their related answers, and achieve a broader view of the problem, all in a single study. Moreover, this method makes it possible to get different perspectives on the problem being researched, by introducing the subjective viewpoint (Ivankova & Wingo, 2018).

Mixed methods can be used for different reasons, such as triangulation, complementarity, interpretation, initiation, problem-solving and so on. These influence how the quantitative and qualitative methods are going to be used in the research. This study comprises and utilises sequential mixed methods for collecting data. This entails using both quantitative and qualitative methods at different stages of data collection. As such, the qualitative results, in this case, will complement and enhance the quantitative findings and will further be used to elaborate and contextualise the data. If at the beginning of the data collection the focus was more on the qualitative data, where the research design was following the structure of an explanatory design, once the data collected from the interviews were analysed, it offered new views that could be further analysed by collecting and analysing quantitative data from Wikipedia. Thus, the mixed-method approach, apart from confirming the results generated by the data collected via the quantitative method with the data collected via the qualitative method, also offered the means to reach complementarity (Saunders et al., 2019, p. 182).

3.4 Qualitative data - Semi-structured interviews

Interviews are the most used method in qualitative research and can be a powerful tool for data collection as they can help in acquiring reliable and valid data which is related to the aim of the research (Bryman 2012, p.469; Saunders et al., 2019, p. 435). The interviews can be influenced by objective or subjective approaches. Employing an objective view when conducting interviews means focusing on finding the needed answer, instead of trying to understand the opinions of the interviewee. A subjective approach seeks to construct meaning by viewing the data gathered as being a product of both interviewees' views and opinions, but also the interviewer's way of responding to the views shared by the participant. Therefore interviews can be divided into structured interviews, which are considered quantitative and qualitative interviews (Saunders et al., 2019, p. 436). Qualitative interviews can be of two types, unstructured interviews or semi-structured interviews (Bryman 2012, p.471). This study employed the semi-structured interview as a method for collecting qualitative data.

Semi-structured interviews are the preferred method as they are flexible, and can provide a thorough amount of data to be further analysed (Bryman 2012, p.473). Hence, these types of interviews can offer in-depth information about the climate change coverage in the English and nordic Wikipedia, which represents the key interest of this research.

Semi-structured interviews come with an interview guide, which contains a list of questions outlined by the researcher. The interviewer, however, can choose to not follow the list of the questions in the exact order specified in the guide, and depending on the answers received, the interviewer can follow up with new questions (Bryman 2012, p.473). Thus, for this research, we created an interview guide that contains the main topics to be covered, although, along the way, new questions will be asked to get a deep understanding of the interviewee's opinions, which can highlight new perspectives.

Even though semi-structured interviews are flexible and allow for a rich amount of data to be collected, they also come with limitations. They are time-consuming, and the process to prepare and set up the interviews, and analyse the data is not an easy task (Adams, 2015; p.66). Additionally, the interview guide can still hold a barrier between the interviewer and interviewee as important aspects of the research can be overlooked if the discussion is not going in a certain way (Bryman 2012, p.473). The interview guide can be seen in Appendix B.

The main method for data collection in this study was the quantitative method utilised, but the semi-structured interviews provided the foundation for a better understanding of the topic, and for digging deeper into the problem. Furthermore, it offered different views and perspectives which we would not be able to perceive only by utilising quantitative methods for the collection of data. The semi-structured interviews were used with the purpose to offer new meaning to the quantitative data and to build the foundation of the research.

The participants in the semi-structured interviews were selected with the use of purposive sampling. This is the most common technique in qualitative research, therefore the sampling was defined and selected by taking into account the research objectives. We tried to select the most relevant participants for this research, as such we looked at Wikipedia's editors which are part of the project WikiProject Climate Change. There are 87 participants currently, however, some are not active in this project, and only a handful of people are still editing. Moreover, we also aimed to find an editor from each language studied, however, that was not possible.

The Snowball sampling technique was also used, as participants in the research proposed other participants to be interviewed. In this research, we interviewed 3 participants of the WikiProject Climate Change.

3.5 Quantitative Data Collection Method

Quantitative research is usually associated with survey research or experiment. In this study, a different approach for collecting quantitative data was selected. This required the manipulation of the data found on Wikipedia, by collecting it via Wikidata API (Application Programming Interfaces). APIs are more and more used as a method for data collection in any type of research (Boegershausen et al., 2022). The well known social media companies' APIs have been of great help to researchers interested in social sciences and humanitarians. The most common APIs in these cases are Twitter API, which allows the extraction of different tweets, Facebook API, Instagram and Spotify (Puschmann & Ausserhofer, 2017).

Wikidata is a free collaborative project that consists of a high number of users and it provides different APIs for connecting to its data (Theo van Veen, 2019; Vrandečić 2014). Wikidata aims to provide a consistent base of knowledge related to its Wikimedia projects, such as Wikipedia, thus its goal is to be central storage for the data coming from Wikipedia and other projects under Wikimedia. Moreover, another objective of Wikidata is to allow other third-party applications and projects to connect to its data. Additionally, it explores the ways through which any individual can query its data (Hernández et al., 2015).

Exactly like Wikipedia, Wikidata is organised into pages and every topic represented on Wikidata is called an entity. Every entity has its own page and they are further divided into items and properties (Erxleben et al., 2014). Each Wikipedia article has a page in WikiData, called items. These items are unique and they can be identified by a name starting with Q, for example, for WikiProject Climate Change the corresponding identifier is Q15305047. To access the data, Wikidata offers SPARQL endpoints, and it can be accessed via the following link https://query.wikidata.org/ (Pillai et al., 2019). Wikidata has a simple data model and makes use of property-value pairs. For instance, the Rome item can have as property *population* and as value the exact number of its population (Erxleben et al., 2014).

Another technique followed in this research for data collection was web scraping, a method used to automatically extract information from any website. Web scraping can be done in two ways, either by manual copy and paste or by connecting to a website's API and extracting the information needed to a database or storing it in a file, this is done via a software program (Gallagher & Beveridge, 2021). This method is

recognised as an efficient way to collect the data available online and it is time-saving, as it allows the rapid collection of high amounts of data. Furthermore, it can be considered a new method for data collection (Khder, 2021).

Web scraping is a method that entails data collection but it is also used for structuring the data so that it is prepared for analysis. Structuring data refers to cleaning up the data to remove unnecessary items so that it can be easily analysed. This method comes with numerous advantages, for instance, it is low cost, efficient and does not require a huge amount of effort from people to perform it (Khder, 2021). One tool used for web scraping in this research was rvest. The rvest package was installed in R studio, which is an integrated development environment for R (Kronthaler & Zöllner, 2021; p.1). With rvest it is possible to read a website simply by making use of the function read_html(). This downloads the HTML of web pages and saves it so that rvest can scrape it. To select different parts of a webpage, rvest comes with another function called html_node, and it requires the URL of the webpage just saved along with a CSS identifier or Xpath. To extract the information from the specified identifier, the following functions can be used, html_attr(), html_table(), and html_text() (Singrodia et al., 2019).

3.6 Ethical Considerations for the research

Ethics are critical in research and should not be overlooked. They should be considered for each stage of research and each data collection method, whether it requires using face to face interviews or the Web to access the data (Saunders et al., 2019, p.232). Thus, the ethical concerns were taken into account for both the qualitative and quantitative methods used.

There are three ethical principles outlined in the Belmont Report which go hand in hand with the ones specified in GDPR (General Data Protection Regulation), and these are r*espect for personas*, *beneficence* and *justice* (Romm, 2018, p.260). Respect for personas requires the researcher to obtain consent from the participants in the study and respect their choices of either taking part or not in the research. Moreover, the consent should be informed, meaning that the participants were given a substantial amount of information regarding the research and how the data will be utilised (Saunders et al., p.266; Romm, 2018, p.366). For the interviews, a form of consent was sent out to the participants which provided further information regarding the aim of the research, and the way their data will be used, and it gave them the option to either participate or not, thus they voluntarily decided to be part of the study. Moreover, the possibility to withdraw from the research was given at any time without having any consequences for the interviewees.

The second principle of the Belmont Report, beneficence refers to causing no harm to the people who are taking part in the research. Harm, in this case, can also refer to the loss of privacy (Barrow et al., 2021). Therefore, the only personal question asked in the interview was related to their background, however, no other private matters were discussed with the participants. Furthermore, the anonymisation and confidentiality of the data is another factor taken into account, thus no personal information will be shared. One participant made it clear that he would like to be mentioned in the research as he is an advocate for this subject, while the rest of the participants will be anonymised. Thus, the GDPR principles of lawfulness, fairness and transparency are respected, as we asked for consent for collecting data, and we were transparent regarding the reasons behind the interview (Voigt & Von dem Bussche, 2017 p.88).

The justice principle is concerned with the fairness in selecting participants to be part of the research. Thus, there was no exclusion based on race, religion, sexual orientation and so on. Excluding persons from the research based on the researcher's prejudices can raise justice concerns, whereas including people who are voluntarily agreeing to participate in the research, and are likely to give their consent is a form of applying the principles of respect and beneficence (Kenneally & Dittrich, 2012). Ethical concerns should also be considered when analysing the data. At this stage it is important that there is no favoured data, no misinterpretation or deception involved (Saunders et al., p272). Little was known about the participants before asking them to be part of the semi-structured interview, as such, no exclusions were done.

When it comes to collecting data via web scrapping from Wikipedia, the ethical principles to follow are related to copyright, and illegal use of data. Web scrapping should not be conducted if it is prohibited in the website's terms of use (Krotov & Silva, 2018). Wikipedia does not forbid web scraping and it allows everyone to share its content Wiki ref. However, the web scraping method should not cause Wikipedia's website any harm. There is a list on Wikipedia with the software that is banned from performing web scraping, and the first ethical consideration in this sense was to make sure to avoid any of the programs showcased on the list. Apart from these, another concern with web scraping is in regards to its purpose. Several laws find illegal the fraudulent use of data collected from web scraping. Moreover, the same principles outlined above in the case of semi-structured interviews should also be taken into consideration if personal data is collected (Krotov & Silva, 2018).

3.7 Analysing the Qualitative Data

Qualitative data analysis is seen as a complex stage in research because text as data is more tedious to reduce and identify different patterns within when compared to using numbers as data (Nowell et al., 2017). According to Thorne (2000), data analysis of

qualitative data as represented in the literature lacks comprehensive discussion. Many papers avoid describing the methods used for data analysis which can lead to confusion among the readers on how the analysis was conducted (Nowell et al., 2017). This paper makes use of thematic analysis to ensure a thorough approach to analysing the data collected via interviews.

Before the analysis of the data that was collected, first, data preparation was conducted. This imposed the transcription of all the interviews recorded which turned out to be an extremely time-consuming process. When conducting the transcription of the interviews, an important factor to take into account is to be sure that a clear distinction can be made between the interviewer and the interviewee (Saunders et al., 2019, p.646). Therefore, the interviewees were identified by a number, for instance, the first participant in the interview is called *Interviewee 1*, whereas, for the interviewer, the identifier used was *Interviewer*, as only one interviewer was present in each interview. For each interview transcripted, a separate file was created which can be found in Appendix C, D and E. Moreover, each file has, in the beginning, a set of keywords which were identified while performing the transcription and were used in the next steps for data analysis.

Thematic analysis is a commonly used method for analysing qualitative data that requires the researcher to identify repeated patterns or themes across data. Thematic analysis can be used for both large and small data sets and it provides a systematic way to analyse the data which can result in comprehensive descriptions of the results (Alhojailan, 2012). Furthermore, this method is flexible as it can be used in any type of study, regardless of the research philosophy or paradigm followed (Saunders et al., 2019, p.651).

This research followed the framework approach which is matrix-based to conduct thematic analysis, and it consists of five interconnected stages*, familiarisation, constructing the thematic framework, indexing and sorting, data summary and display, and mapping and interpretation* (Hackett & Strickland, 2018).

The familiarisation stage started the moment the interviews were transcribed, as the process required continuous listening to the interviews and writing every piece of information in the transcription. Moreover, the transcripts were read and reread along with the analysis so that the data was better understood and recurring themes could be identified across the data set (Saunders et al., 2019, p.651). Additionally, keywords were identified for each interview transcribed to easily understand the patterns that can be formed around the data (Hackett & Strickland, 2018).

In the second stage, constructing the thematic framework, the keywords were used to identify the top-level categories of the framework. Three main categories

emerged, Wikipedia Structure, Wikipedia Contributors and WikiProject Climate Change, which addressed the questions related to Wikipedia's governance, meaning how Wikipedia operates, questions related to its contributors and the reasons behind their choice of voluntarily editing the online encyclopedia, and questions addressing the WikiProject Climate Change, its aim, challenges, participants and solutions to its level of coverage across the other Wikipedias.

During stage three, each paragraph in the transcription was inserted into an excel along with more keywords identified after another round of actively reading the transcripts. Additionally, the interview was coded and themes were identified and assigned to each segment of the transcript. For each interview, a thematic analysis file was created which includes the number iterations that was needed for coding the interviews and creating the code book which can be seen in Appendix F, H and H. *Themes* for this research are defined as categories identified in the research by finding repeated patterns or important statements. The process of recognizing different categories emerging from the data set followed an inductive approach. Thus, these themes do not necessarily reflect the questions asked to the participants, as they are resulting from the data available for analysis. This method was used as it can offer a wider and more comprehensive analysis of the data (Kiger & Varpio, 2020). Each interview required between two and three iterations of both the code book and interview in order to reach a final thematic analysis.

For the data summary and display stage, the data of the interview was matched with the top-level categories as well as the sub-themes and concepts that were further identified. During the last stage, the data was mapped with the codebook resulting in the final framework of the data analysis., The final overview of the codebook contains the top categories, sub-themes, concepts, their overall description and their corresponding statements taken from all the interviews and this is showcased in Appendix I. The final framework was used to present the results related to the climate change coverage in the English edition compared to the nordic editions of Wikipedia.

3.8 Analysing the Quantitative Data

The quantitative data was collected and analysed by using R with rStudio. R can be explained as a program that offers a different set of tools for the researcher to perform diverse methods for data analysis. R provides a wide range of packages that can be used to either clean up the data or visualise it and plot it with the use of graphs. rStudio is an application in which all the functions that R provides can be used as it comes with a user-friendly interface for the researchers to use when analysing the data (Kronthaler & Zöllner, 2021; p.1).

Using the package rvest available in R, we scraped Wikipedia and gathered the information needed for this research. Two main pages were harvested, first the article page *List of Wikipedias* and the second, the WikiProject Climate Change. Once we managed to select the information from different HTML attributes from Wikipedia, for instance, *table,* we conducted exploratory data analysis (EDA) to first understand the structure of the data. Here the library tidyverse was used as it comes with many packages that helped to organise and visualise the data.

The list of Wikipedias returned a data frame containing 326 rows, which was further divided into one subset containing only the necessary information for the English and Nordic Wikipedia. The WikiProject Climate Change offered us different lists that were converted into data frames to be able to visualise the data by creating graphs with the use of the ggplot2 package available with the tidyverse library. Using only rvest and web scraping as a method for data collection provided data in regards to general information about the languages and also about the WikiProject Climate Change on the English Wikipedia.

To collect the number of articles under the WikiProject Climate Change on the English Wikipedia, SPARQL was used. The query was run under Wikidata's query editor and the results were exported and imported into R for further analysis. The same steps were done to query how many of the currently available articles on the WikiProject are also found in the Nordic editions. The package *pageview*s was utilised to gather the articles with the most page views related to climate change so that we can understand what articles are viewed the most, and what sparks the interest of the readers. All the data collected was visualised with ggplot2 and the results of the analysis are further presented in section 4.

3.9 Reliability and Validity in this research

The quality of the research is an important matter that should not be overlooked. Any research and findings should be open to be assessed in terms of accuracy of its results, the methods used for data collection and the final assumptions reached. Without a clear understanding of the findings, research can be seen as futile (Long & Johnson, 2000). The most common criteria for evaluating the research are reliability and validity (Bryman, 2012, p.46).

Reliability is concerned with replication and consistency. If a researcher is able to replicate a study and achieve the same conclusions and results, then that research is deemed as reliable (Saunders et al., 2019, p.214). While reliability is important to ensure the quality of the research, it is not enough to only base the quality of your study on it. The quality of any research is assessed by both reliability and validity (Saunders et al.,

2019, p.215). Validity can be understood as the integrity of the findings and it is assessed by looking at how the instrument used for data collection measure what it was supposed to measure (Long & Johnson, 2000). These two terms are mostly used for quantitative studies, however their meaning was altered so that they can be applied to qualitative studies as well (Bryman, 2012, p.389).

To ensure the reliability of the semi-structured interviews conducted in this study, a few criteria were used. Previous researchers suggested that to evaluate the quality of the qualitative studies, reliability and validity should not be used, but instead, other concepts should be applied, which represent just an alternative to reliability and validity and these are trustworthiness and authenticity (Bryman, 2012, p.390). The principles of *trustworthiness*, such as *credibility, dependability, transferability* and *confirmability* were used in this research to assess the quality of the semi-structured interview.

Credibility parallels internal validity, and it encapsulate the fact that the research findings are credible, and it follows an accurate analysis of the data by highlighting the original views of the participants (Anney, 2014). A way through which credibility can be demonstrated is by utilising respondent validation as technique. This technique has different forms and the researcher can use the one deemed appropriate (Bryman, 2012, p.391). In this study, the interviewees received a copy of the transcripts, so that they can provide validation on what was said during the interviews. Moreover, having two different methods for data collection, quantitative and qualitative depending on how they were used can impose the concept of triangulation. A method of triangulation was used in this study, as the purpose of the interviews was to provide a deeper understanding to the results outlined by the quantitative method, to make sure that the data is more thorough and vigorous.

Transferability is an alternative to external validity and is assessing whether the findings can be generalised to other contexts (Anney, 2014). Since qualitative methods usually have a small sample of subjects having certain characteristics in common, the findings generated are unique to that specific context (Bryman, 2012, p.392). However, according to Bitsch (2005), a qualitative researcher can showcase the transferability context by having thick descriptions and purposive sampling. Thus, this study offered a thorough description of the methodology, both from a data collection and data analysis perspectives. Moreover, purposive sampling was used to select participants in the study to maximise the collection of data that is of interest to this research.

Confirmability is a term used to replace objectivity and it highlights that even though complete objectivity in qualitative research is not feasible, the findings are clearly formed by analysing the data and are not presenting the view of the researcher, but the view of the participants (Bryman, 2012, p.393). Confirmability goes hand in hand with

dependability, thus, since in this study, the transcripts were documented and saved, as well as the steps in which thematic analysis was conducted, it can be concluded that confirmability was also reached (Anney, 2014).

Dependability is parallel to reliability and it requires having records of all phases of the research so that if needed they can be reviewed (Bryman, 2012, p.391). To ensure high dependability of the semi-structured interview, an interview guide was created containing the questions to be asked, their descriptions and reasoning for asking them. Moreover, the interviews were all transcribed, each transcription having its own file and they were further sent for verification to the interviewees. The analysis of the interview, by following the framework approach, requires extensive and comprehensive descriptions of all iterations. Thus, dependability can be achieved by following an auditing approach to data collection and analysis. Additionally, intercoder reliability was also applied. For this, an extra coder was needed, as the interviews were analysed just by one person initially. The extra coder reviewed and coded the last iteration of the interviews and Cohen's Kappa was calculated. Cohen's Kappa is the most used reliability measure to address the issue of random agreement (Warrens, 2015). To understand the coefficient, Landis and Koch (1977) introduced the following guideline: 0.00 - 0.20 show slight agreement, 0.21 - 0.40 means fair agreement, 0.41- 0.60 moderate agreement and 0.61 - 0.80 substantial agreement, while 0.81 -1.00 high agreement (Warrens, 2015). The Cohen's Kappa coefficient for the performed interviews is as follows: 0.61 in the case of top-level categories and 0.40 in regards to sub-categories. The Cohen's Kappa calculation can be found in Appendix J..

When it comes to the quantitative data collection techniques used, web scraping and APIs, a solution to prove validity is to keep the raw data generated by the two methods. The source code of Wikipedia was saved in R as well as the files resulting from SPARQL, together with the data collected by connecting to Wikipedia's API (Boegershausen et al., 2022). One thing that is worth mentioning is that the web data is not static, thus since the point of collection is different from the current date, if compared to the present the data may be different. If Wikipedia's editors tagged more articles as part of the WikiProject Climate Change from the point of the data collection that data is not reflected in this study. Moreover, sometimes the different languages used for programming the user interface of different websites can hide elements that the web scraper would miss (Boegershausen et al., 2022; Xu et al., 2020). However, this did not represent an issue in this paper due to the simple user interface of Wikipedia.

4. Results

In this section, the findings of both methods applied in this study are presented. This includes the results derived from the quantitative data collection and RrStudio analysis, as well as the results coming from the qualitative data collection and thematic analysis. Regarding quantitative results, first, the general findings in regards to Wikipedia will be presented, followed by the ones strictly describing the climate change coverage in the English and Nordic editions.

4.1 Interview Results

The thematic analysis of the interviews brought to light the following findings.

All the interviewees are participants of the WikiProject Climate Change and they all have a rich experience when it comes to editing on Wikipedia. One of the interviewees, however, apart from being an editor on Wikipedia, is also the lead strategist for the Wikimedia Foundation. Thus, he shared not only the view coming from the knowledge related to being an editor on Wikipedia, but also the point of view coming from his experience working for the Wikimedia Foundation.

The interviews were coded into three main categories, called top-level categories. The first category is Wikipedia Structure which contains all the statements related to Wikipedia in general and its governance. Second, Wikipedia Contributors, here all the information regarding the editors on Wikipedia, such as their experience and the challenges they had to face are included. And third, WikiProject Climate Change, which encapsulates all the comments regarding the WikiProject Climate Change, challenges and solutions to the information coverage in other languages.

The top-level category Wikipedia Structure has as the most repeated themes, Wikipedia's governance and Wikipedia editing practice, both having 11 repetitions. Followed by Wikipedia Climate Change Coverage which was identified as being the theme related to the climate change subject and its coverage on the non-English Wikipedias and it is part of the Wikipedia Structure top-level category, having 7 repetitions. When it comes to the WikiProject Climate Change top-level category, this has both WikiProject Climate Change governance as one of the themes appearing the most, 9 times, and WikiProject Climate Change Coverage theme which is related to the information strictly related to climate change knowledge on Wikipedia but concerning the WikiProject Climate Change.

Wikipedia Structure has a sub-category called WikiProjects which describes the way a WikiProject works in general. Even if it can be deduced that it is the same as WikiProject Climate Change top-level category, that is not the case as the latter represents only the information strictly related to the project Climate Change on Wikipedia. Therefore, the interview results presented are related to the themes appearing the most. The themes and sub-themes identified during analysis, as well as their number of repetitions, can be seen in Appendix K.

WikiProjects are usually formed on the bigger Wikipedias languages edition because they have a larger number of editors and they are a response to the wish of the contributors interested in finding other editors to work together with:

> "Well, so I think Wiki projects are like a manifestation of a desire that anyone working on knowledge wants, which is like, who are my peers?" (Appendix C, min 07.33)

WikiProjects are not formed in a structured way, they are impulsively decided upon, and since Wikipedia was built, no one has checked on how different the infrastructure of one WikiProject is from another in the English Edition. They work so well on the larger language editions due to the set of tools and automation that were set up. However, a WikiProject would not be maintained at the same level in other languages, as the tools used in one edition are not localised. (Appendix C, min 23:38)

The WikiProject Climate Change was formed between five and ten years ago, under a different name, WikiProject Environment, however, it was forgotten up until a small group of editors decided to bring it back to life as they realised the importance this topic had. The articles included in the project are decided by the editors and the number of articles currently showcased on Wikipedia, 3758, is believed to increase as not all of the articles related to climate change have been assessed because this is still considered a new WikiProject. As all interviewees mentioned, the project represents a way of bringing together the editors who are interested in improving climate change content:

> "Basically, a Wiki project is a group of contributors who want to work together to improve Wikipedia. So it's just a way of organizing the group's efforts at creating and improving articles." (Appendix E, min 01:59)

> "A group of us, kind of Phoebe Mires, myself, and Sue Lane, and Femke, and Elizabeth all kind of showed back up and the Wiki project, because [...] this was like a forgotten structure. It has some like real, real needs, right? It needs to be nurtured and needs to be like, thought about." (Appendix C, min 07:33)

> “The Wiki project, it's a couple of years old, it used to be a working group or whatever it's called under WikiProject Environment. But when climate change started to dominate the news on its own, right, we decided to make its own Wiki project.”(Appendix D, min 02:03)

> “Then you find the research communities that you want to go to right? You join spaces where people are having the same discussion you want to have, right? That is like at the heart of knowledge production. This is how knowledge production works in the world right now. On Wikipedia, Wiki projects are a response to that, they are a response to that impulse.” (Appendix C, min 27:37)

One of the interviewees stated that some editors prefer to work alone, while most the people with topical expertise would be more interested to collaborate with other editors to improve the articles which they are all passionate about. Editors collaborate via the WikiProject Climate Change and by utilising the talk pages of each article (Appendix C, min 53:38)

The information asymmetry in regards to climate change and its coverage on Wikipedia is a well-known issue to Wikimedia Foundation’s lead strategist. As a matter of fact, he became an editor simply to be able to provide his skills to the community of practice being formed around the climate change topic and to understand how to build a structure around them. Additionally, he points out that the information asymmetry is a result of a narrative issue and not a knowledge problem:

> “And I, I kind of like I'd been seeing the knowledge gap around climate change. But it was like, became absolutely clear that like, this could be a very powerful community of practice, and the movement and so I, in order to be able to facilitate that community practice, I also needed to learn how to contribute to it right? And so I learned how to edit on these topics. And it's actually largely a narrative problem, not a not a knowledge problem.”(Appendix C, min:1.48)

The narrative problem raised by the lead strategist is also highlighted by another interviewee, which puts into the light the fact that an article on a smaller language Wikipedia is unlikely to be updated as often as articles on the English Wikipedia do, thus articles have may contain information which is not relevant anymore (Appendix D, min: 04.19). Moreover, the editors, after one article published by BBC came out which showed that there exists a bias in the smaller language editions of Wikipedia, have worked on updating the other editions as well. The bias came from an outdatedness in the content as climate change is a topic that evolves and with every year more and more information is shared:

> “So this is a translation of the English Wikipedia in 2007, into 2007 there was I think I'm not entirely sure. But I think there was a big enough minority of scientists that believed the sun caused up to one third or one fourth of climate change, to wear and inclusion. And that viewpoint is now fringe. And it's a really small minority of scientists. So it doesn't deserve to be in the article anymore. But I think in 2007, it was fine. Because non-English, Wikipedia does get updated way less. Once the article is there, it doesn't get touched anymore. And that false balance continued to be there 15 years since. So that was the case in more than half of the articles we reviewed.” (Appendix D, min 04:49)

Regarding the coverage of the climate change topic in other languages, all the interviewees are aware of the asymmetry. One mentioned that the top articles, which are the most important articles according to Wikipedia’s method of grading are part of the other editions (Appendix E, min 14:19). The interviewees, however, have split opinions when it comes to the level of coverage of climate change articles in the Nordic editions in comparison to the English Wikipedia. One of the interviewees stated that a study should be conducted where the way Danish, Swedish and Norwegian people read Wikipedia is investigated. As it takes a lot of work to improve these editions, this effort would be meaningless if they do not read in their local language, but instead, choose the English Wikipedia:

> “Just that I sometimes I don't know how much these other language, languages get used. [...] I don't know if anyone has done a study to really look at which people look for information on the Swedish Wikipedia and which people, which Swedish people go straight to the English Wikipedia? So one would have to weigh up the effort. If it's a lot of effort to improve it in the other languages, but if the other languages don't get used that much, then yeah, is it worth the effort? Especially given that people can use Google Translate anyway? You can. It's getting more and more accurate. So I just don't know, in the longer term, whether it's worth it to have humans translate things when you've got to the Translate, which is pretty good, too.” (Appendix E, min 05:17)

Whereas, another interviewee believes that by having different levels of coverage between the editions is the way through which misinformation happens in the online encyclopedia. Additionally, it was also pointed out that that the other editions require more manual work than the English Wikipedia in order to be maintained (Appendix C, 38:55). And even if one wants to semi-automate these languages, it will require a lot of work, as they will not to understand the entire ecosystem and Wiki markup for that specific language (Appendix C, 23:38). The communities formed on the

Nordic Wikipedias are not large enough to ensure that their articles are up to date. Even the English Wikipedia is facing a challenge in this regard, as its fewer read articles are having issues with outdatedness, which is a result of missing more active editors (Appendix D, min 20:56).

The views of the articles can be an incentive for improving certain articles, at least according to one of the participants in the interview. Some editors choose to focus on the articles that are read the most, but it was also pointed out that some articles' views will increase once their content is improved:

> "So if, if an article has high page views that I think it's worth putting the effort in, but sometimes it can be related. So sometimes, the page views will go up once the once the article is better."(Appendix E, min 13:16)

In regards to the quality assessment process on Wikipedia, the articles to reach feature articles or even good articles go through a very demanding process of peer review (Appendix C, min 46:08). Moreover, the article needs to source all the up to date sources and it needs to be extensively covered, for climate change, the article should have between 8000 and 10000 words to be deemed as comprehensive. Having articles of higher quality was highlighted as an objective of the WikiProject Climate Change:

> "Yeah, as in, I'm pushing for the objective, because I think it's really important for articles to be reviewed, and to get as many eyes as possible, as that ensures both, neutrality and verification accuracy." (Appendix D, min 16:18)

This assessment of the quality of different articles serves as a backlog for the editors. Feature articles mean that those articles are accurate and they trust that the topic is comprehensively covered. The other grades are used to understand what needs more work and where to focus so that the article can be assessed and regraded (Appendix C, min 46:08).

Another interesting fact stated which is related to Wikipedia's structure, more precisely its governance was that editors on other language editions will revert changes made by other editors if they do not share the view they supported:

> “Sometimes? Yeah. Most of the projects were really happy that we did this. Some of them weren't. Belarusian Wikipedia was really like no, we want to keep this false balance in, because we believe this is neutrality to give both types of people a voice. The there was one where the I think it was? Was it Serbian? Croatian? No, it wasn't. There was one language where two admins disagreed strongly about what should be in there. And in the end, are the sort of the non-climate denial admin one out there like: please don't those start another argument between the admins here because there was already so much tension!” (Appendix D, min 25:56)

As the knowledge gaps and information asymmetry issue is a matter known by the Wikimedia Foundation, they are running different campaigns to try and get new editors as part of Wikipedia in response to the fact that active editors are nearly not enough. They are also working on designing new interfaces and tools for editors to easily find actions that they can do. For instance, the editors can introduce a topic of interest and then they will see a list of articles related to that specific subject which they can further work on. However, as this is machine automated tool, it still needs improvement to make sure it displays the right related articles (Appendix C, min 27:37).

A solution for improving the coverage of climate change articles in the Nordic editions was to have semi-automated topical communities on these Wikipedias. Another solution was to have students involved in editing Wikipedia as part of their English classes, where for example the Swedish students could look over the content on the local Wikipedia and compare it to the English edition. It only takes a bit of curiosity and fine storytelling techniques to be able to make improvements on Wikipedia. Additionally, Wikipedia is a subject that should be included in our normal lives, and if people work in the field of climate change communication they should maybe consider editing the content of the online encyclopedia as part of their job:

> “And then if it could be brought in more into people's normal jobs, so if if you work for a company that does anything on climate change communication, then you maybe should spend time in your normal job, like four hours a week or one hour per month or whatever to improve that so it should be, shouldn't just be seen as a volunteer thing where people do it as a hobby on the weekend. Yeah, I think it'd be more integrated into people's day jobs. “ (Appendix E, min 28:20)

The interview findings shaped how the quantitative analysis will be conducted as different points that could be interesting for investigation were addressed. It was stated that the active editors on Wikipedia are not enough on the English edition, let alone on the other languages encyclopedia. Therefore, the first part of the quantitative analysis was to look into the general issue about Wikipedia, for instance, the number of editors in the English compared to Nordic editions, the number of active editors and the total number of articles. Another statement that influenced an investigation when conducting the quantitative analysis was to see if there is any correlation between the page views of the articles and their improvements. Additionally, the interviewees believed that the top articles, the most important articles in the WikiProject Climate Change can be found in the other editions, as such they will be a focus of this study analysis.

4.2 Quantitative Results

This research aims to investigate climate change coverage on Wikipedia with a special focus on the English edition and Nordic countries, which were defined as being Denmark, Sweden and Norway. However, the first information we were interested in collecting was the general information about these editions, such as the number of articles, number of users, and number of active users, which can highlight from the beginning if there is any notable asymmetry between the above mentioned Wikipedias.

In total, Wikipedia has in the English Edition 43,429,630 contributors, followed by the Swedish edition with 817,600 editors. The only other language that has above 500,000 editors, respectively 549,737 is the Norwegian (Bokmål) language. In contrast, Norwegian (Nynorsk) has the lowest number of contributors among these languages, with a total of 119,325, while the Danish Wikipedia has 427,448 editors. This can be seen in figure 4.

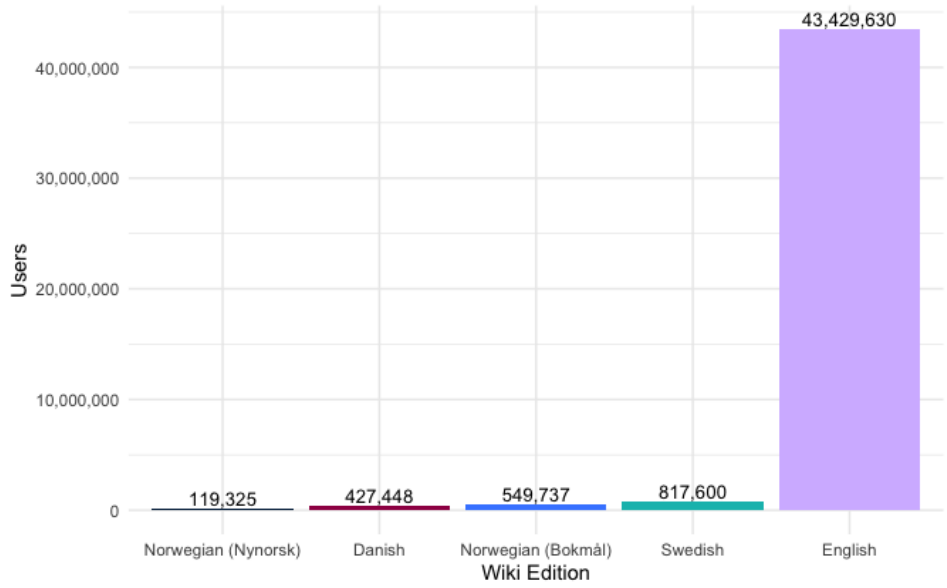


Figure 4. The number of editors for the English and Nordic Wikipedia

Even though it seems that Wikipedia has a substantial amount of users, especially English Wikipedia, only less than 1% of them are active. In fact, 99.7% of all

users are inactive on the English Wikipedia, as well as on the rest of the Nordic editions. The total number of active users can be seen in figure 5.

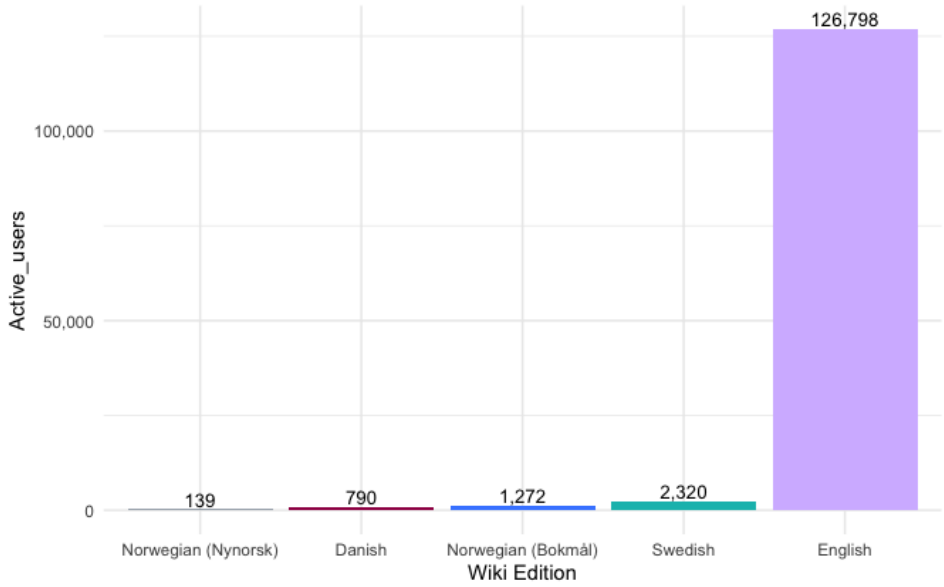


Figure 5. Total number of active users on the English and Nordic Wikipedia

When it comes to the number of articles, English Wikipedia is again the one with the highest total, reaching 6,487,057 articles. The Swedish edition has 60% fewer articles, having a total of 2,566,791 written articles, whereas the Norwegian (Bokmål) has 90% fewer articles than the English Wikipedia. The Danish Wikipedia has a total of 274,765 articles, while the Norwegian (Nynorsk), the smallest Wikipedia in this sample, registers 161,792 articles, which represents a drop of 97% compared to the English edition. This can be seen in figure 6.

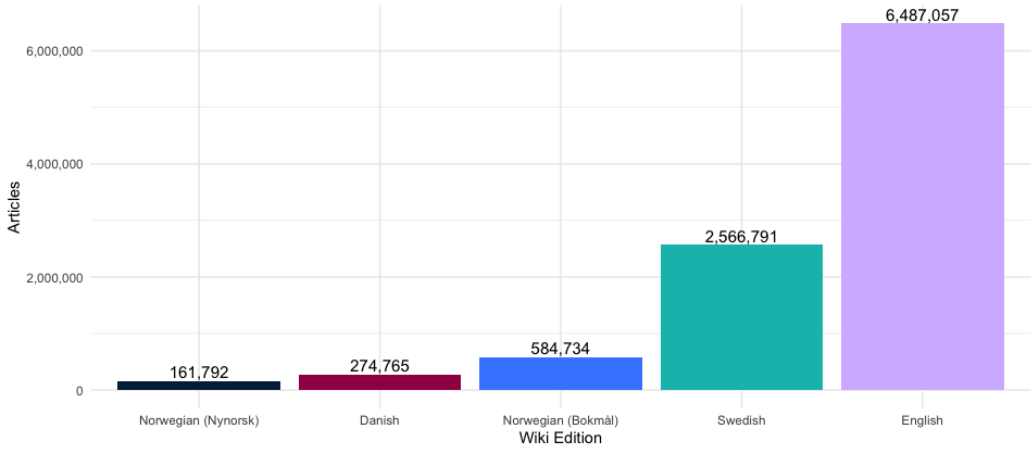


Figure 6. Total number of articles on the English and Nordic Wikipedia

A notable difference can also be seen in the number of edits registered in the English and Nordic Wikipedias. As it is showcased in figure 7, the English edition currently has above 1 billion edits, whereas the Swedish Wikipedia has 50,297,072 total edits done. The difference between these two editions is 95%. The Norwegian (Bokmål) edition has over 22 million edits, while the Norwegian (Nynorsk) registers only 3,356,043 million edits. The Danish Wikipedia has a total of 11,091,218 edits.

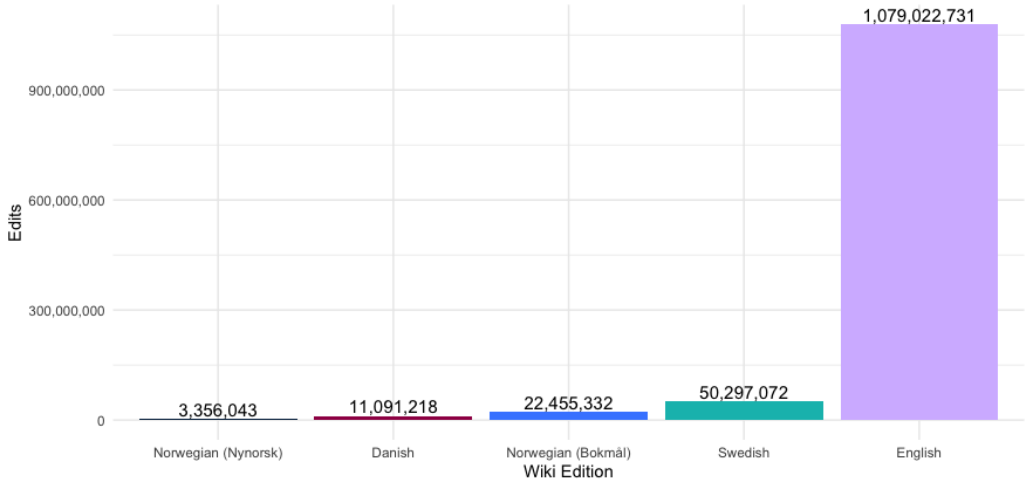


Figure 7. Number of edits in the English and Nordic Wikipedia

It can be concluded from these findings concerning the general information on Wikipedia that there exists information asymmetry between the Nordic editions and the English Wikipedia. There is a significant difference between the number of articles available in the English edition compared to the Nordic countries' editions. Additionally, there is a gap between the number of editors and active editors found on these Wikipedias.

4.2.1 WikiProject Climate Change results

The first interesting finding regarding the WikiProject Climate Change was done by simply checking the project page and looking at which language editions the project is linking to. Among the Nordic countries, only the Swedish Wikipedia has a page about the WikiProject and it does not have the same outline nor the same content as it can be found in the English edition. The WikiProject Climate Change on the English Wikipedia has a more thorough description, which includes the top edited articles in the project, as well as all the task forces that were created within WikiProject Climate Change and its participants. In contrast, the Swedish page of WikiProject Climate Change has a paragraph on how a user can contribute to this project, a summary of its participants, and a short backlog of the tasks that should be carried out in this project. The Danish and

Norwegian (Bomkål), as well as the Norwegian (Nynorsk), do not have a page for the WikiProject Climate Change.

The WikiProject Climate Change has in the English edition 3758 articles and this information can be found on Wikipedia. However, when the collection of data happened, only 3527 articles were found by using SPARQL. Upon deeper investigation, the missing 200 articles were not extracted in the first place because some of the articles were drafts which were not approved to be an article yet on Wikipedia, others were templates, while others were paragraphs related to climate change in different articles. But as Wikipedia's data is considered to be the master data and trustworthy, the missing 200 articles were also included in the data analysis. As such, as it can be seen in figure 8, the English edition has 3758 articles, whereas the other languages all have below 500 articles. This means that more than 87% of the articles found in the English edition are missing from the Nordic countries' editions.

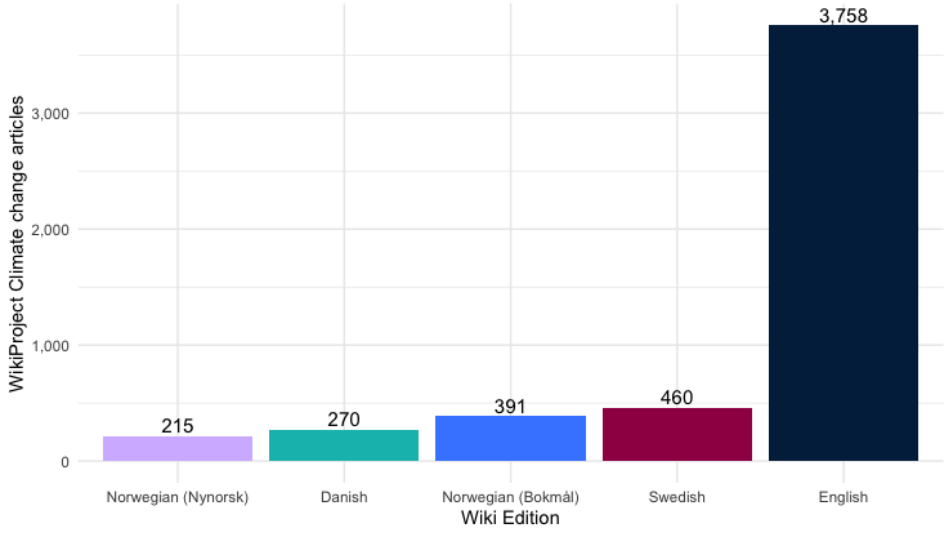


Figure 8. WikiProject Climate Change articles on the English and Nordic Wikipedia

On Wikipedia, articles can be graded by their quality and importance, as such, we were also interested to get an overview of the quality of climate change articles in both the English and Nordic editions. In figure 9, it is presented the quality and importance of the WikiProject Climate Change in the English edition. The importance is

showcased in the columns, and it can be from top which is the highest, to low, which is the lowest importance, however, the table also presents the number of articles of which the quality and importance are not defined. The grades of the articles are represented on the row level and they can be from FA, feature articles and highest quality to stub, lowest quality.

Quality and Importance of Climate Change Articles in the English Wikipedia

Quality	Top	High	Mid	Low	Not_Defined	Total
FA	2	1	2	2	0	7
GA	2	7	19	13	1	42
B	10	66	73	68	0	217
C	43	197	247	588	47	1122
Start	5	96	237	778	300	1416
Stub	1	5	44	298	191	539
List	1	11	9	14	5	40

Figure 9. Importance and Quality of WikiProject Climate Change articles in the English Wikipedia

As it is highlighted in the figure 9, the English Wikipedia has a total of 7 FA, which is the highest grade, out of which 2 have the highest importance. Moreover, it has a total of 42 GA, and just like in the feature article case only 2 are considered to be of high importance. Additionally, it has a total of 217 B-class articles, which means that these articles reasonably cover the topic and they have a good structure. The highest number of climate change articles in the WikiProject is part of the Start quality, which are the articles that do not have a comprehensive description of the topic and they lack citations.

Regarding the Nordic countries' editions and their quality, the results which were extracted with SPARQL are shown in figure 10. These articles do not have an importance assigned, therefore they could not be analysed. Additionally, with SPARQL,

you can only look to see if the articles have a Wikidata badge, which can help identify their quality, and only feature articles and good articles have a badge assigned.

Quality of Climate Change Articles in the Nordic editions

Language	Feature_Article	Good_Article	Feature_List
Norwegian (Bokmål)	9	7	1
Swedish	0	3	0
Danish	0	1	0
Norwegian (Nynorsk)	2	0	0

Figure 10. Quality of the climate change articles in the Nordic editions

As it can be seen in figure 10, out of 391 articles, the Norwegian (Bokmål) Wikipedia has a total of 9 feature articles, 7 good articles and one feature list, while the Swedish edition has only 3 good articles from a total of 460. The Danish Wikipedia has only one good article, and the Norwegian (Nynorsk) edition has 2 feature articles.

Since the SPARQL query returned a list containing only feature articles and good articles, as these are the only ones that could be found, a deeper analysis was conducted. The interview results highlighted that the top articles, the articles defined to have the highest importance in the English Wikipedia should be present in the other language editions. Therefore, to further investigate the quality of the articles on the Nordic Wikipedia, the top articles having the highest three quality grades such as FA, GA and B-class were selected. This provided us with a list of 14 articles for analysis.

In figure 11, the results are presented. Out of the 14 articles found in the English Wikipedia, 6 were not found in the Norwegian (Bokmål) edition, 5 articles were also missing from the Swedish Wikipedia, whereas 4 articles were missing from the Danish edition and 8 articles were absent from the Norwegian (Nynorsk) Wikipedia. In both Norwegian (Bokmål) and (Nynorsk) editions, 6 articles were not graded, while 9 articles were also missing quality assessment on the Swedish and Danish Wikipedia. A total of 4 articles were assessed in terms of their quality, out of the most important articles analysed. Hence, 3 were FA in Norwegian (Bokmål) Wikipedia, and one B-class article in the Danish edition.

Quality of Climate Change articles in the Nordic editions

Language	Total_Feature_Articles	Total_Good_Articles	Total_B_Class_Articles	Total_Articles_Without_Grade	Total_Missing_Articles
Norwegian (Bokmål)	3	0	0	6	6
Svedish	0	0	0	9	5
Danish	0	0	1	9	4
Norwegian (Nynorsk)	0	0	0	6	8

Figure 11. Top Articles Quality in the Nordic Wikipedia

Besides from investigating the quality of the climate change articles which are part of the WikiProject Climate Change, we were also curious to study what people read, and which articles spark the interest of the readers. The time selected for the investigation was from 2015 up until 2022 on a monthly basis.

Figure 12 showcases the list of all articles found to be part of the top 1000 viewed articles in the online encyclopedia and the total number of times it was a part of this list during the specified time frame. The *Global Warming* article appears 31 times, followed by the *David Attenborough article,* which was found in the list a total of 18 times. The article about *Greta Thunberg* was part of the top 1000 most viewed articles on Wikipedia 15 times from 2015 to 2022, whereas the *Al Gore* article appeared 5 times, followed by the *Human overpopulation* article which appears 4 times.

Top most viewed articles in the English Wikipedia

	article	Count
14	Global warming	31
10	David Attenborough	18
17	Greta Thunberg	15
5	Al Gore	5
18	Human overpopulation	4
11	Don't Look Up	3
12	El Niño	3
16	Green New Deal	2
19	Hurricane Ida	2
21	Keystone Pipeline	2
23	Rice	2
24	Sustainable energy	2
1	2015 United Nations Climate Change Conference	1
2	2019–20 Australian bushfire season	1
3	2020 California wildfires	1
4	2021 United Nations Climate Change Conference	1
6	Antarctica	1
7	Bushfires in Australia	1
8	Clathrate gun hypothesis	1
9	Climate change	1
13	Extinction Rebellion	1
15	Great Reset	1
20	Insulate Britain protests	1
22	Paris Agreement	1
25	World energy consumption	1

Figure 12. Top viewed Climate Change articles on the English Wikipedia

Out of these articles, the top 10 most viewed articles in regards to climate change are highlighted in figure 13, showcased below. The highest rank was reached by the article about Greta Thunberg, which had rank 6, having more than 3 million views. However, this article is also the third and fourth on the list, reaching high news 3 times in 2019 in 3 months. the Paris Agreement article in 2017 was the 16th most viewed article

on Wikipedia, with more than 2 million views. Sustainable Energy is the 10th most viewed article in our list, however, the highest rank it reached was 179 back in January 2020.

Top 10 most viewed articles

project	language	article	access	granularity	date	rank	views
wikipedia	en	Greta Thunberg	all-access	month	2019-09-01	6	3209862
wikipedia	en	Paris Agreement	all-access	month	2017-06-01	16	2243752
wikipedia	en	Greta Thunberg	all-access	month	2019-12-01	48	1488575
wikipedia	en	Greta Thunberg	all-access	month	2019-10-01	53	1190911
wikipedia	en	David Attenborough	all-access	month	2020-10-01	75	1137775
wikipedia	en	Bushfires in Australia	all-access	month	2020-01-01	78	1231506
wikipedia	en	2019–20 Australian bushfire season	all-access	month	2020-01-01	79	1215629
wikipedia	en	Keystone Pipeline	all-access	month	2021-01-01	95	1179737
wikipedia	en	Al Gore	all-access	month	2020-11-01	127	1064958
wikipedia	en	Sustainable energy	all-access	month	2020-01-01	179	831480

Figure 13. Top 10 most viewed articles

To get a better understanding of the views over the period of time selected, the articles which appeared more than 3 times in the top 1000 viewed articles were selected and analysed with time series plots.

Figure 14 highlights the Global Warming articles views. The article started to register a substantial increase in the views from August 2019, when it increased from less than 350.000 views to more than 425.000 views in one month and reached the highest views in November 2019. Figure 15 showcases the article about David Attenborough and its views. It can be seen an increase in the page views from September 2020 to October the same year, when it reached its peak of more than 1 million views. However, after October 2020 its views drastically decreased.

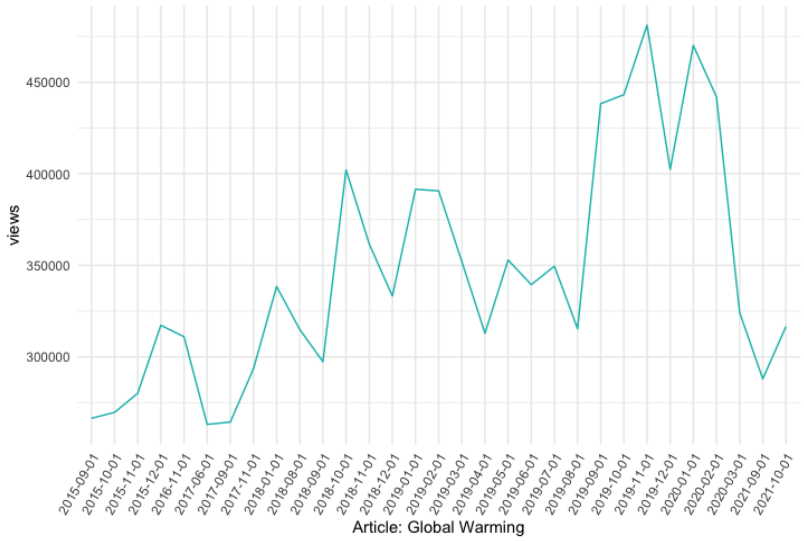


Figure 14. Global warming views

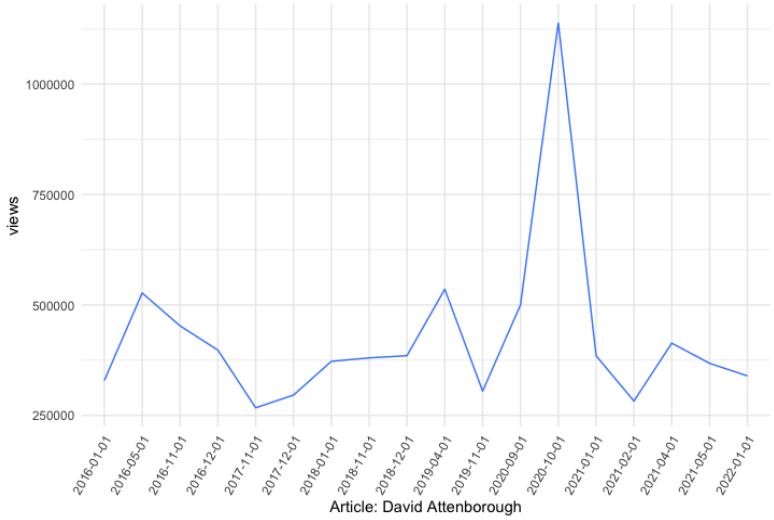


Figure 15. David Attenborough article views from 2015 to 2022 in the English Wikipedia

In figure 16, it can be seen that the article regarding Greta Thunberg, the young climate activist, registered more than 3 million views in September 2019, however, after that the views decreased substantially in one month.

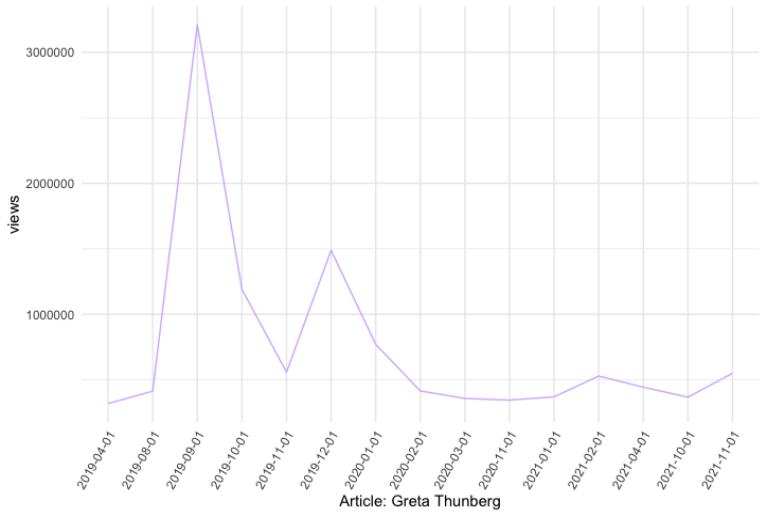


Figure 16. Greta Thunberg article views on the English Wikipedia.

The Al Gore article is also among the ones which were part of the top 1000 viewed articles more than 3 times, and its views are presented in figure 17. It can be seen that the article had the most views in November 2020, however, after that this article did not make it into the top viewed articles list. In figure 18, the last article's views during our selected time are shown. Just like in the previous case, Human Overpopulation reached the highest views in March 2018, but after that moment it did not make it back into the top 1000 most viewed articles.

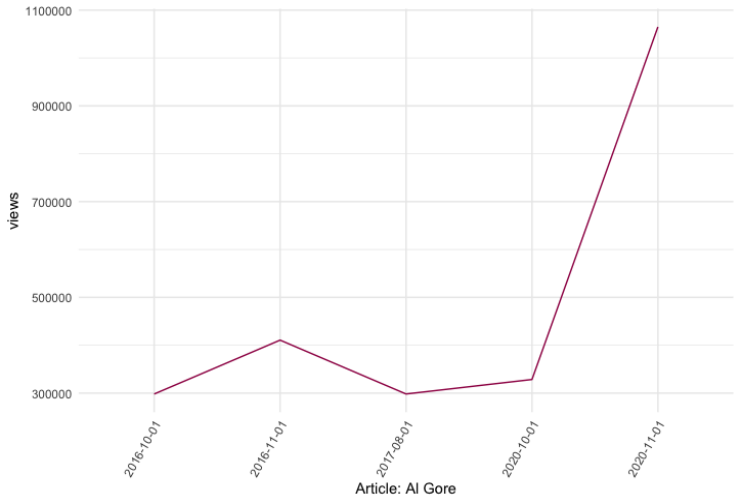


Figure 17. Al Gore article views in the English Wikipedia

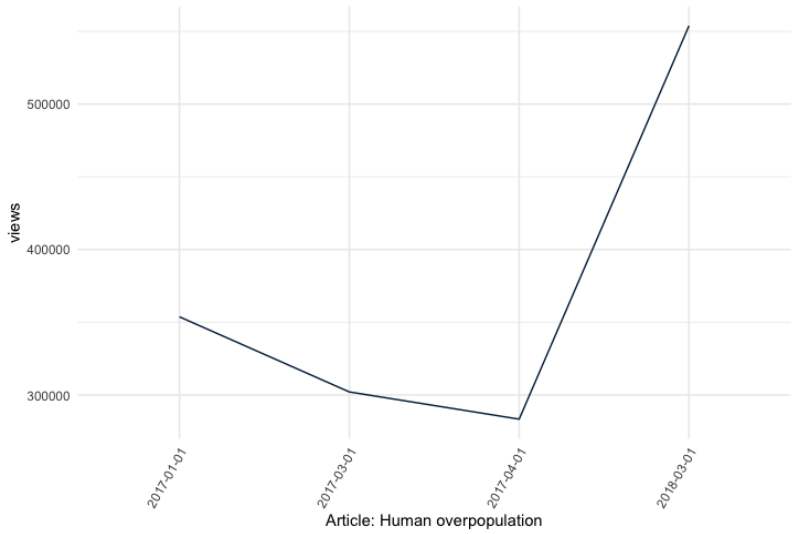


Figure 18. Human overpopulation article views in the English Wikipedia.

We have shown the views of the climate change articles in the English Wikipedia. When it comes to the Nordic countries, their top most viewed articles which are part of the WikiProject Climate Change are quite similar to the results showcased in the English Wikipedia.

Figure 19 highlights the most viewed articles in Norwegian (Bokmål) Wikipedia, and the number of times they appeared in the top 1000 most viewed articles, this is showcased under the column count, as well as their ranks and views. This edition has just like in the English Wikipedia, articles such as Global warming (Global opvarming), Greta Thunberg, El Niño, Paris Agreement (Parisavtalen), and Antarctica in the top 1000 most viewed articles. However, there are new articles on the top such as Sustainability (Bækreft), Greenhouse gas effect (Drivhuseffekt) and Renewable energy (Fornybar energi). Regarding their views and ranks, the article about Greta Thunberg is the most viewed reaching a rank of 6, followed by David Attenborough reaching 118 as rank, and in third place is the article about the Paris Agreement with a rank of 155.

Top most viewed articles in the Norwegian (Bokmål) Wikipedia

	article	Count
7	Global oppvarming	27
8	Greta Thunberg	27
1	Antarktis	17
3	David Attenborough	14
6	Fornybar energi	12
4	Drivhuseffekt	10
9	Parisavtalen	2
2	Bærekraft	1
5	El Niño	1

Most viewed articles in the Norwegian (Bokmål) Wikipedia and their views and ranks

article	date	rank	views
Greta Thunberg	2019-09-01	6	54772
David Attenborough	2016-05-01	118	8849
Parisavtalen	2017-06-01	155	6172
Global oppvarming	2019-04-01	306	5141
Drivhuseffekt	2016-04-01	334	5116
Antarktis	2016-11-01	377	5425
Fornybar energi	2016-05-01	414	4473
El Niño	2015-12-01	675	3418
Bærekraft	2020-10-01	766	3049

Figure 19. Top most viewed articles in Norwegian (Bokmål) Wikipedia, their ranks and views

Regarding Norwegian (Nynorsk), its top viewed articles, their ranks and views are highlighted below, in figure 20. Unlike English and Norwegian (Bokmål), the top viewed article, in this case, is the one about Greenhouse gas, reaching rank 53 and a total of 968 views. Followed by one article which was not present before in the other languages, Climate in Asia (Klima i Asia), which is also the one appearing more often in the top most viewed articles. The third most viewed article is Antarctica which reached rank 185. 2 other new articles are worth mentioning that were not found in the other

languages, and these are Climate in Antarctica (Klima i Anarktis) and Köppen climate classification (Köppen si klimaklassifering).

Top most viewed articles in the Norwegian (Nynorsk) Wikipedia

	article	Count
6	Klima i Asia	58
2	Drivhusgass	54
3	Fossil energi	18
1	Antarktis	16
5	Klima i Antarktis	2
7	Köppen si klimaklassifisering	2
4	Greta Thunberg	1

Most viewed articles in the Norwegian (Nynorsk) Wikipedia and their views and ranks

article	date	rank	views
Drivhusgass	2015-11-01	53	968
Klima i Asia	2016-05-01	144	528
Antarktis	2021-02-01	185	396
Fossil energi	2015-11-01	203	392
Klima i Asia	2020-01-01	203	317
Greta Thunberg	2019-09-01	465	180
Köppen si klimaklassifisering	2018-11-01	468	277
Klima i Antarktis	2017-01-01	777	181

Figure 20. Top viewed articles in Norwegian (Nynorsk) edition, their ranks and views

Similar to both the English and Norwegian (Bokmål) edition the Danish Wikipedia has as the highest viewed article the one about Greta Thunberg. Sustainability (Bæredygtihed) is the second-highest viewed with a rank of 231, followed by David Attenborough article. Their views and ranks can be seen in figure 21.

Top most viewed articles in the Danish Wikipedia

	article	Count
7	Global opvarmning	29
2	Bæredygtighed	24
8	Greta Thunberg	24
1	Antarktis	19
10	Vedvarende energi	12
6	Fossilt brændstof og brændsel	9
3	David Attenborough	7
4	Drivhuseffekt	7
5	Drivhusgas	1
9	Klimaendring	1

Most viewed articles in the Danish Wikipedia and their views and ranks

article	date	rank	views
Greta Thunberg	2019-08-01	7	40732
Bæredygtighed	2016-09-01	231	5149
David Attenborough	2018-08-01	250	4666
Antarktis	2017-01-01	321	5680
Drivhuseffekt	2017-06-01	424	3583
Global opvarmning	2017-06-01	429	3569
Bæredygtighed	2017-06-01	665	2822
Vedvarende energi	2016-09-01	665	2961
Fossilt brændstof og brændsel	2016-11-01	691	3481
Klimaendring	2019-10-01	699	139
Drivhusgas	2017-06-01	939	2346

Figure 21. Top most viewed articles in the Danish edition, their ranks and views

The Swedish Wikipedia, as is highlighted in figure 22, similar to the other countries has as the most viewed article the one in regards to Greta Thunberg, which is

also the one who is seen the most often in the top 1000 viewed articles in this edition. A new article found on this Wikipedia edition among the top most viewed article is the *2015 United nations climate conference in Paris*.

Top most viewed articles in the Swedish Wikipedia

	article	Count
8	Greta Thunberg	39
2	Antarktis	31
9	Växthuseffekten	23
7	Global uppvärmning	6
4	El Niño	5
3	David Attenborough	3
1	Al Gore	2
5	Förenta nationernas klimatkonferens i Paris 2015	2
6	Förnybara energikällor	1

Most viewed articles in the Swedish Wikipedia and their views and ranks

article	date	rank	views
Greta Thunberg	2019-09-01	7	111587
David Attenborough	2016-05-01	99	19675
Greta Thunberg	2020-04-01	274	13948
Växthuseffekten	2016-05-01	274	12155
Antarktis	2016-12-01	476	10480
Al Gore	2020-11-01	477	11571
Förenta nationernas klimatkonferens i Paris 2015	2015-12-01	555	9314
Global uppvärmning	2019-05-01	575	9171
Förnybara energikällor	2016-05-01	879	6963

Figure 22. Top most viewed articles in the Swedish Edition, their ranks and views

Based on the findings from the interview results, that mentioned that sometimes if an article's quality is improved their views might increase, we were interested to see if there exists a correlation between an article's quality and its page views. As such, just like in the case with the quality of articles in Wikipedia, the same list of articles was selected as they were found to be the most important articles in the project. As it can be seen in figure 23, the Climate Change article sees a decrease, in the views after the improvement, whereas the article about Paris Agreement is seeing an increase from less than 50,000 to almost 100,000 views.

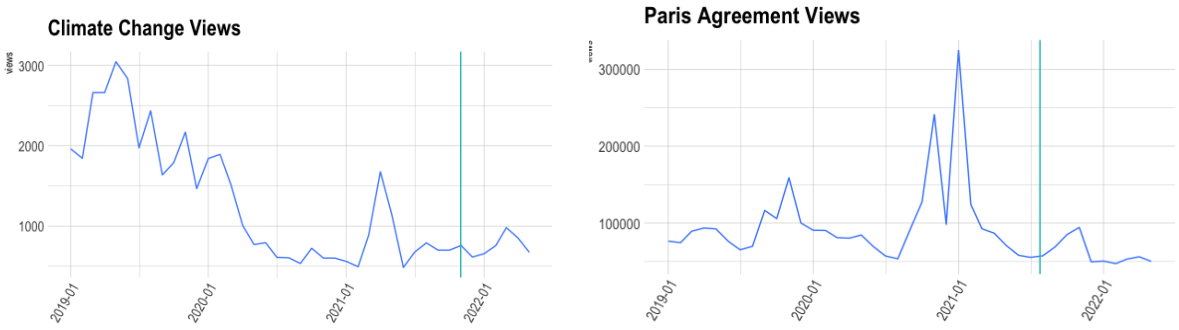


Figure 23. Climate Change and Paris Agreement views before and after quality improvement

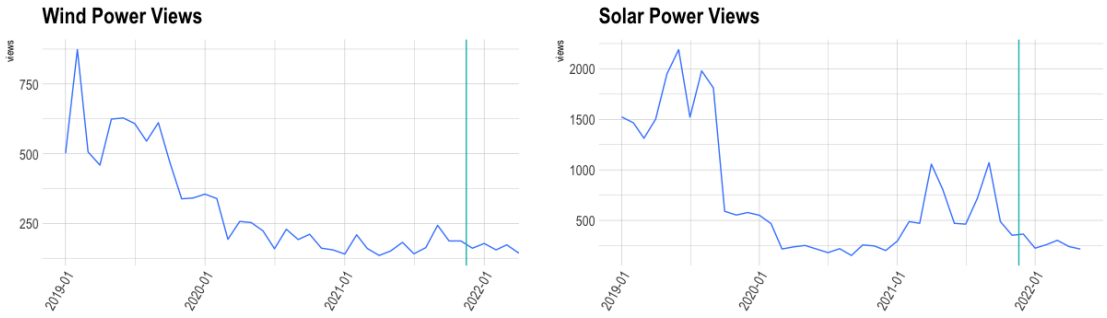


Figure 24. Wind and Solar Power articles views before and after improvement

When it comes to the articles about Wind and Solar power, as it is showcased in figure 24, their views are decreasing after their quality was improved.

In the case of the Renewable Energy article, for a couple of months after the improvement of the article the views increased, to suddenly drop in July 2020. However, after that point in time, the views noticed an increase again. Whereas for the Sea Level Rise article, its views slightly decreased right after improvement, increased again reaching its peak at more than 500.000 views. However, since mid-2020, it did not have such a substantial increase, see figure 25.

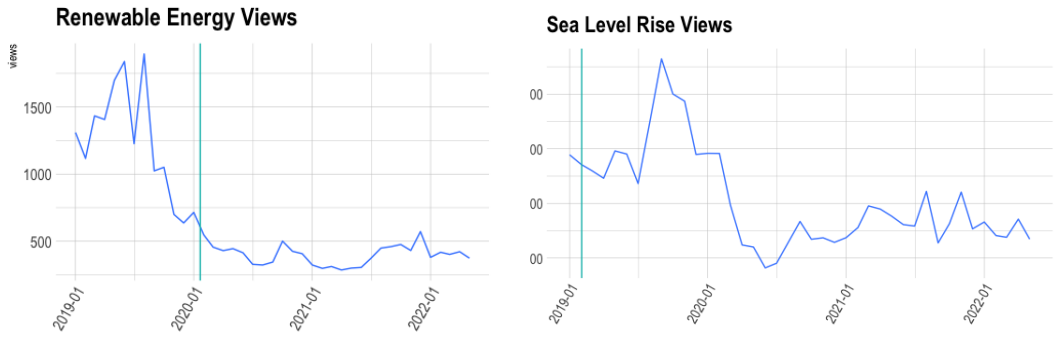


Figure 25. Sea Level Rise and Renewable Energy articles' views before and after improvement.

The article about renewable energy was also found in the Danish and Norwegian (Bokmål) editions. Therefore, their time-series graphs were also analysed.

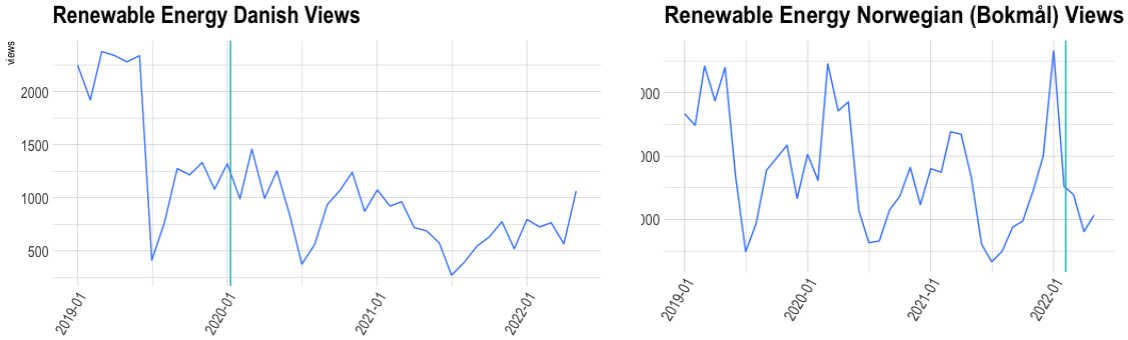


Figure 26. Renewable Energy views of Danish and Norwegian (Bokmål) edition

Figure 26 highlight the views on both Danish, on the left, and Norwegian (Bokmål) on the right. For Norwegian (Bokmål) Wikipedia, the page views of the article did not increase but decreased. Until now, no substantial increase in the views can be highlighted. For the Danish edition, the views decreased right after improvement for one month, to slowly increase in March 2020 and reach a peak of 1458 views. However, they registered a substantial decrease in mid-2020, and up until now, they did not reach 1458 views again. However, the peak of 1458 views cannot be related to the improvement of the article.

Regarding the Climate Change article, just like in previous cases, the page views decreased after improvement in the quality, to steadily increase after. This is highlighted in figure 27.

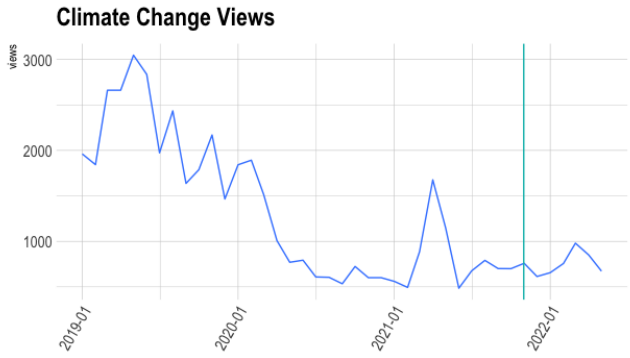


Figure 27. Climate Change views

The Climate Change article was found as a feature article in Norwegian (Bokmål) and its views are presented below, in figure 28. After improvement for one month the views did not increase nor did they decrease. However, two months later there can be seen a sudden and substantial increase in the page views.

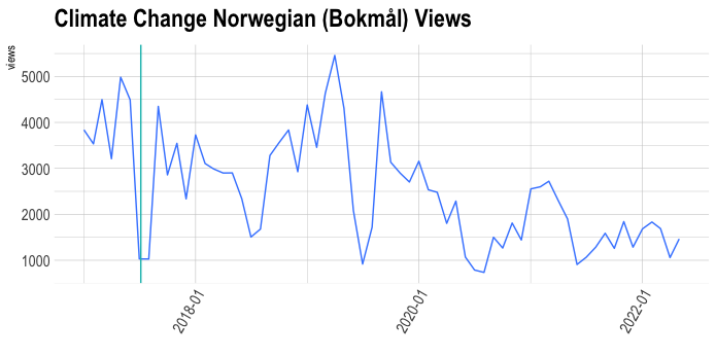


Figure 28. Climate Change views in the Norwegian (Bokmål) Wikipedia

In the case of the Sustainable Energy article, the views did not increase and this can be seen in figure 29.

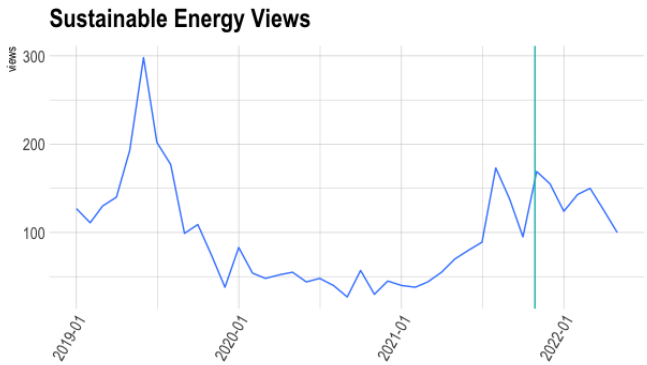


Figure 29. Sustainable Energy views

The last two articles which were identified as the most important, Greenhouse gas and Effects of Climate Change are represented in figure 30. For Effects of Climate Change, the number of views stayed the same for almost one year, whereas for the Greenhouse gas article the views increased. However, there was already a tendency of an increase in the views as it showcased, therefore it cannot be said that the increase is due to the quality improvement of the articles.

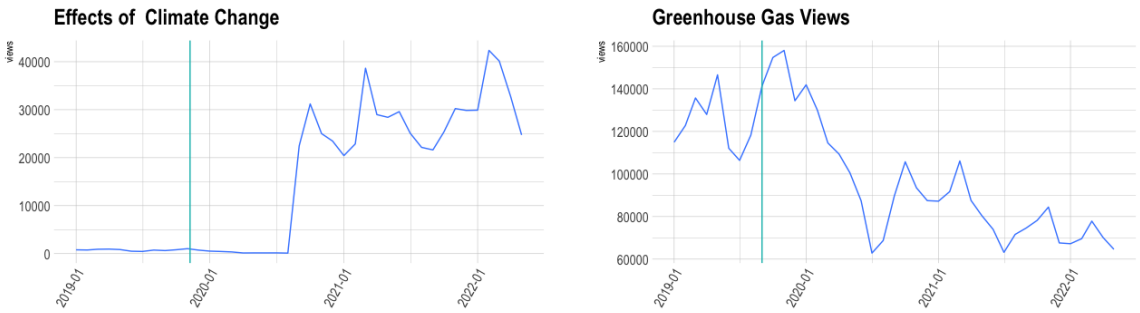


Figure 30. Effects of Climate Change and Greenhouse gas article views

The Effects of Climate Change article was also found on the Norwegian (Bokmål) Wikipedia with a current assessment of FA. Therefore its views can be seen in figure 31.

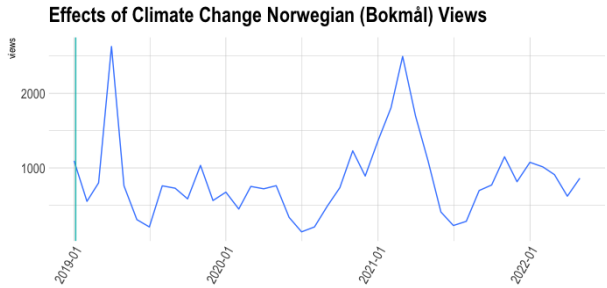


Figure 31. Effects of Climate Change vires in the Norwegian (Bokmål) Wikipedia

As figure 31 shows, the article was improved at the beginning of January 2019, but its views decreased after the assessment of the article.

Based on the quantitative analysis results, from the beginning, there was a notable difference in the content of the English Wikipedia compared to the Nordic editions, as the English encyclopedia has more editors and a wider number of articles.

When it comes to the WikiProject Climate Change, there is an information asymmetry between the Nordic countries and the English Wikipedia. More than 87% of the articles which were found in the English edition were absent from the Nordic countries' Wikipedia. Moreover, a deeper analysis of the most important articles in regards to the WikiProject Climate Change, showed that the Nordic editions were missing some of these articles. Norwegian (Nynorsk) had a total of 8 articles that were not found, out of the 14 analysed articles, while Danish had 4 missing articles, Swedish 5 and Norwegian (Bokmål) 6.

Contrary to the interview results that showed that there could be a connection between the quality of an article and its page views, the quantitative analysis showed differently. Most of the articles noticed a decrease in the views, others did not register any change in the first month after their quality improved. Some articles after a couple of months did register an increase, but it cannot be related to their quality being better.

5. Discussion and Limitations

This section presents the discussion in regards to the findings collected via semi-structured interviews and quantitative method such as web data logs, in order to compare the literature already existing in the field of study. This section will be structured in three subsections, each one related and following the order of the research questions outlined in the problem field.

5.1 Wikipedia and its solutions to knowledge gaps

In this subsection, the results attained for the first research question are presented. The first research question was:

How can Wikipedia make sure that climate change subject is covered at the same level in the English and Nordic Wikipedia?

As it was highlighted in the interview results, the Wikimedia Foundation is in fact aware of its knowledge gaps, also in regards to the topic of climate change. The lead strategist mentioned that as one steps away from the English encyclopedia, the other editions are more manual and they do not have the same set of tools to maintain certain projects. Additionally, it was pointed out that the English Wikipedia does not have enough active editors to make sure that the content is accurate, with the English edition being the one having the highest amount of active editors when compared to the Nordic Wikipedia. According to the lead strategist, who is also a contributor to the encyclopedia, there is only a handful of editors who are active on the WikiProject Climate Change. Thus, it can be concluded that Wikipedia needs more editors to make sure that its content stays accurate. This is supported by the work of Schneider et al. (2014), which stated that ongoing editing and new editors are important for making sure that Wikipedia will continue growing and will be maintained. Moreover, Morgan et al. (2013), showed that Wikipedia should focus on making sure that the newcomers are aware of the community that stands behind Wikipedia, and that having more diversity among the contributors will increase the chance of the new editors finding similar like-minded people and feel integrated, which can lead to their retention.

According to the interview results, different campaigns and events have been organised by Wikimedia Foundation to attract new editors. And a solution proposed to solve the knowledge gaps between the English and Nordic Wikipedia was to have semi-automated communities for these editions. To make it easier for its community of

editors, the English Wikipedia has numerous tools and bots some of which can be used to detect the edits which do not respect Wikipedia's guidelines. This is also mentioned in the work of Faulkner et al. (2012) which shows that a semi-automated community can increase the productivity of the new editors (Faulkner et al., 2012). But for new contributors to be able to contribute in a high-quality manner to Wikipedia, they need to learn fast the norms around Wikipedia. Failing to onboard the new users properly can in fact lead to frustration and even isolation. Thus, another solution suitable for gaining new editors and making sure Wikipedia is growing and maintained is suggested by Narayan et al., (2017). Hereby a gamified solution to the onboarding process of Wikipedia is proposed.

Another proposal outlined in the interview results was to involve high school students in Wikipedia. The teachers could utilise the language classes, and for example, Sweden can have students to compare the English articles to the ones found in their local edition and improve it. This goes hand in hand with the WikiEd project proposed by the Wikimedia Foundation as a response to the need for new contributors. This program urge university teachers to have Wikipedia writing assignments as part of their curriculum (Li et al., 2020). It was shown to be a successful method of attracting new long term editors and improving the quality and coverage on Wikipedia, as the students edited more than 24,000 articles (Li et al., 2020).

This research aimed to study the coverage and quality of climate change topics on Wikipedia and find a solution to the gaps identified. The information asymmetry is caused by the differences in the amount of active editors that it are present between the editions. We found that the English Wikipedia has 126,798 active editors while the Swedish has 2,320, Norwegian (Bokmål) 1,272 and Norwegian (Nynorsk) 139 editors. Therefore, attracting, retaining and integrating new editors by Wikipedia is necessary to sustain these editions. Having a program such as WikiEd on the Norwegian, Danish and Swedish editions could have the same results as it did on the English Wikipedia and it could be a step further in closing the knowledge gap about climate change in Wikipedia.

5.2 Wikipedia and Climate Change Coverage

The discussion in regards to the climate change coverage in the English Wikipedia compared to the Nordic Editions related to the second research question is outlined:

RQ2: How many articles about climate change are in the English Wikipedia compared to the Nordic language editions and what is their quality?

Our results showcase that from the beginning when investigating the general information about Wikipedia, such as the number of articles, editors and active editors,

we could see a difference between Nordic Wikipedia compared to the English edition. The Swedish Wikipedia has 60% fewer articles than the English edition, while the rest of the Nordic countries have a difference of at least 90% when compared to the English Wikipedia. Redi et al., (2020) described a wide range of knowledge gaps that are affecting Wikipedia, such as contributor, reader and content gaps. Among the reader gaps, the language was highlighted. Depending on the language that an individual can read, it will have different consequences in regards to the content that is available to them and can impose certain obstacles if they have to read in a language that they are not familiar with (Redi et al., 2020).

The WikiProject Climate Change has a total of 3,758 articles, however, the Nordic editions all have less than 500 articles, highlighting a discrepancy of more than 87%. One of the interviewees mentioned that the top articles, which are the most important articles according to Wikipedia's article importance grading, are found in the other languages. However, out of the 14 selected articles for the analysis, 8 were not found in the Norwegian (Nynorsk) edition, 4 on the Danish Wikipedia, 5 of the articles were not covered in the Swedish, and 6 were missing in the Norwegian (Bokmål) edition. The substantial difference in the content between the languages in the encyclopedia was also pointed out by Roy et al. (2020), as the non-English content available in the other languages was found to be a small fraction of the English Wikipedia. Moreover, our findings highlight an asymmetry in the content between the Nordic editions, with Swedish Wikipedia being the largest among the four, followed by Norwegian (Bokmål), Danish and lastly Norwegian (Nynorsk). The significant difference between the Swedish and the Danish edition was also showcased in the work of Bick (2014).

The disrepancy between the English and the Nordic edition was also found in regards to the quality of Wikipedia. In the English Wikipedia, articles tend to have a grade, such as a feature article, good article, B-class and so on. However, even though the grading system is also part of the Nordic editions, it is not applied to all of the articles. The analysis was conducted on the 14 top articles selected, and it showed that both Danish and Swedish had more than half not graded, more precisely 9 articles, while both Norwegian (Bokmål) and (Nynorsk) had 6 not assessed. A total of 7 FA, 42 GA and 10 B-class were found in the English edition. The Danish and Swedish Wikipedia both had 0 FA articles, and 1 GA for Danish, while the Swedish Wikipedia had 3 GA. Surprisingly, the Norwegian (bokmål) edition had 9 FA, and more feature articles than the English Wikipedia. Wikimedia Foundation urge the improvement in the coverage of important topics that impact our world and that can potentially improve people's life (Redi et al., 2020). However, our results show that a topic that represents the challenge

of the contemporary world is approximately 13% covered in Nordic editions when compared to the English Wikipedia.

The purpose of investigating how the articles which are part of the WikiProject Climate Change are covered in the English and Nordic editions was to discover what type of knowledge gaps can be highlighted. This can highlight different ways that Wikimedia Foundation can use to bridge the gaps and improve the coverage of climate change.

5.3 Climate Change Articles and readers' interest

RQ:3 How often do climate change articles get visited and what sparks more interest in the readers?

The findings in regards to the page views of the climate change articles on the English and Nordic Wikipedia showed that the readers are interested in almost the same articles in all editions. Greta Thunberg, El Niño, Antarctica and Climate Change (Global opvarming, Klima ændring) articles are among the top viewed articles in all languages. Articles such as Renewable Energy (Fornybar Energi - Norwegian (Bokmål), Vedvarende Energi - Danish), Greenhouse Gas (Drivhusgas), and Greenhouse Gas Effect () were also found among the top viewed, but only in the Nordic editions.

Looking closely at the article about Greta Thunberg and its corresponding page views, the same date was noticed in all languages, respectively 1st September 2019, which corresponds with the time when the youth strike for climate change was happening (Boulianne et al., 2020). On the English Wikipedia, we also noticed articles like Paris Agreement, and Australia Bushfires being among the top viewed articles. Just like in the case of Greta Thunberg article, they were visited around the exact time when these events were occurring. According to Singer et al. (2017), Wikipedia is mostly used for quick information, and readers are influenced by the media to read different articles in the encyclopedia. 30% of the respondents in the survey conducted by Singer et al, (2017) said that they are reading Wikipedia due to the media. Hara & Doney, (2015) showed that there exists a correlation between events occurring and the number of edits an article received. Editors on Wikipedia showed high interest and they put a lot of focus on the topics which are related to breaking news. This was also highlighted in the interview results, where the editors started a cross Wiki project to eliminate the bias found in regards to climate change after a BBC article came out describing the current situation in Wikipedia (Silva, 2021). Moreover, another study showed that Reddit posts influence Wikipedia's page views (Moyer et al., 2015).

The interview results highlighted that the articles' views might increase once their quality is enhanced, for instance being deemed as FA. Additionally, it was also underlined that some editors prefer to focus on and improve the quality of the articles which have higher views. Our quantitative results did not prove that the page views of an article are related to the fact that its quality was improved, as only 2 articles noticed a slight increase, while the rest saw a decrease. Warncke-Wang et al. (2015), showed that subjects which are in high demand by the readers, do not have high-quality coverage on Wikipedia. On the contrary, topics that were in low demand were the ones having the higher quality. Therefore, being aware of the articles which are of interest to Wikipedia's readers can help the editors to make a certain decision about which articles to be improved. However, even if certain articles are not in high demand, but if they are regarding important topics they should be thoroughly covered (Redi et al., 2020).

The page views can tell a lot about Wikipedia and in this case, it showed us the most visited articles from the WikiProject Climate Change in the English and Nordic Wikipedia. It seems that the interest of the readers in specific topics is correlated with breaking news or certain events occurring in the world. However, we also saw that currently the topic of climate change is not among the top 1000 viewed articles on Wikipedia.

5.4 Future Work and Limitations

This research studied the climate change coverage and its quality in Wikipedia, with a focus on the English and Nordic editions. The work presented in this paper is not without limitations. First, limitations were imposed by the method for the data collection, the semi-structured interviews. Even though they are flexible, and the interviewer can ask follow up questions, since only one interviewer was present, it can imply that other views may have been overlooked to investigate.

Second, we interviewed participants of the WikiProject Climate Change. However, a possibility could exist that certain contributors edit articles related to the WikiProject Climate Change even though they are not part of the project. Moreover, our study focused on the Nordic countries as well and having the perspective of the editors from those communities would have given different views to the solution proposed by this study.

From the interview results, a point that could be a further investigation of the research was the readership of the Nordic Wikipedia. Thus, looking into the page views for the specific project would show how many people read the specific language and where the traffic is coming from. This could reveal if Nordic countries read the English Wikipedia more, where the climate change topic is wider covered, or their local

languages editions, where there is a substantial discrepancy in the coverage and quality of the articles related to climate change.

Future work should also seek to investigate on a deeper level the quality of the articles in the Nordic editions. The SPARQL query returned only the FA and GA articles. We did look into the articles which were classified as top articles in the English Wikipedia for the FA, GA, and B-class. However, it would be beneficial to understand how the rest of the articles are classified. As they can highlight where the editors in the Nordic community should focus on when it comes to improving the quality of climate change articles in their edition.

Moreover, some solutions were proposed, for instance, WikiEd, and even though they were studied in the past and had great results, they were investigated in the context of the English Wikipedia. However, to be able to present these further, a study should be conducted where the WikiEd would be applied to the Nordic countries. Therefore experiments or use cases could be utilised where one school from each country is selected and different hypotheses could be tested regarding the possible outcomes that could arise from the experiments.

6. Conclusion

Climate Change is a topic of high importance as it is considered the challenge of the century. The science behind climate change is becoming more certain and there is a wide consensus that climate change will have major impacts on our society.

Wikipedia, part of Wikimedia Foundation, was shown to be a good communication channel for scientific topics such as climate change, as it is one of the most popular websites nowadays. The movement behind Wikimedia set a goal that by 2030 Wikimedia will reach knowledge equity. Currently, Wikimedia Foundation is working to find solutions for bridging the different knowledge gaps identified across Wikipedia. Although new editors are a necessity for sustaining Wikipedia, the results presented in this research highlighted that even the English edition is missing more active editors. Moreover, this could be noticed in the Nordic Wikipedia, where the Swedish version had the highest amount of contributors, and content, whereas Norwegian (Nynorsk) was seen to be the one having the least contributors and articles. Furthermore, the climate change topic is not presented at the same level in the Nordic countries, compared to the English edition. A substantial amount of the articles that can be found concerning this subject in the English Wikipedia are missing from the Danish, Norwegian and Swedish editions. Moreover, among the articles found, few of them were rated as GA or FA, while the rest of the articles were missing an assessment. As the mitigation policies for climate change depend on public awareness, and Wikipedia being

one of the main sources of information used for retrieving knowledge about climate change, the gaps in the content could lead to important facts being omitted regarding this topic and its mitigation.

This research also highlighted the possible connection between the media or other events occurring sparking an increase Wikipedia's page views. The article about Greta Thunberg was visited the most when the youth strike for climate change happened. Moreover, the Australian Bushfires article was also visited the most when the catastrophic event occurred back in the beginning of 2020. Wikipedia is connected with other websites as well, for example, one of them being Reddit, and posts that were done about the encyclopedia lead to an increase in page views. This can be powerful information to be understood in the field of climate change communication as maybe new channels or platforms could be used to make the climate change information available to a wider audience.

With this study, we also tried to identify different approaches that Wikimedia Foundation could take to solve the climate change information asymmetry discovered. The possible solutions identified goes hand in hand with the literature studied. A good example is the English Wikipedia and one that proved to have the best results in the past was WikiEd project. This could be made available for the Nordic languages. The purpose hereby is to have teachers include in their curriculum Wikipedia writing assignments. Moreover, it was pointed out that the WikiProjects participants use different sets of tools to make their work easier and understand which articles need to be assessed. However, the small languages require more manual work. Thus, having these tools available in other languages could potentially be a solution to the problem of climate change coverage, as it can help editors in their day to day tasks.

Wikipedia is considered an important source of knowledge and in this study, we showed that there is information asymmetry regarding climate change coverage in the English Wikipedia and the Nordic countries. As climate change is a topic that gained a lot of attention lately, of which mitigation and adaptation depend on the people's awareness of climate change, it is necessary to offer the right information to all the individuals, no matter which language they speak.

References

Adams, W. C. (2015). Conducting semi-structured interviews. *Handbook of Practical Program Evaluation, 4*, 492-505.

Alhojailan, M. I. (2012). Thematic analysis: A critical review of its process and evaluation. *West East Journal of Social Sciences, 1*(1), 39-47.

Anderson, A. (2011). Sources, media, and modes of climate change communication: the role of celebrities. *Wiley Interdisciplinary Reviews: Climate Change, 2*(4), 535-546.

Anney, V. N. (2014). Ensuring the quality of the findings of qualitative research: Looking at trustworthiness criteria. *Journal of Emerging Trends in Educational Research and Policy Studies, 5*(2), 272-281.

Ayers, P., Matthews, C., & Yates, B. (2008a). *How Wikipedia works: And how you can be a part of it*. No Starch Press.

Baker, E., McNamara, L., Mackay, B., & Vincent, K. (2020). How to Contribute Climate Change Information to Wikipedia. 7, 8, 13, 16, 18, 22

Barrón-Cedeno, A., Paramita, M. L., Clough, P., & Rosso, P. (2014a). A comparison of approaches for measuring cross-lingual similarity of wikipedia articles. Paper presented at the *European Conference on Information Retrieval,* 424-429.

Barrow, J. M., Brannan, G. D., & Khandhar, P. B. (2021). Research ethics. *StatPearls [Internet]* (). StatPearls Publishing.

Bassani, E., & Viviani, M. (2019). Automatically assessing the quality of Wikipedia contents. Paper presented at the *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing,* 804-807.

Bick, E. (2014). Translating the Swedish Wikipedia into Danish. Paper presented at the *Swedish Language Technology Conference, , 2014*

Bitsch, V. (2005). Qualitative research: A grounded theory example and evaluation criteria. *Journal of Agribusiness, 23*(345-2016-15096), 75-91.

Blennow, K., & Persson, J. (2009a). Climate change: Motivation for taking measure to adapt. *Global Environmental Change, 19*(1), 100-104.

Boegershausen, J., Datta, H., Borah, A., & Stephen, A. (2022). Fields of Gold: Web Scraping and APIs for Impactful Marketing Insights.

Boulianne, S., Lalancette, M., & Ilkiw, D. (2020). “School strike 4 climate”: social media and the international youth protest on climate change. *Media and Communication, 8*(2), 208-218.

Brügger, A., Morton, T. A., & Dessai, S. (2015). Hand in hand: Public endorsement of climate change mitigation and adaptation. *PloS One, 10*(4), e0124843.

Bryman, A., & Bryman, A. (2012). Social research methods (4th ed.). Oxford University Press. 38, 46, 59, 199,389, 390, 391, 392, 393, 469, 471, 473, 501, 628, 670

Budžytė, A., & Balžekienė, A. (2018). PUBLIC PERCEPTIONS OF INSTITUTIONAL RESPONSIBILITY IN CLIMATE CHANGE RISK IN BALTIC NORDIC COUNTRIES. *Journal of Security & Sustainability Issues, 7*(4)

Cook, J. (2022). Understanding and countering misinformation about climate change. *Research Anthology on Environmental and Societal Impacts of Climate Change,* , 1633-1658.

Cooke, A. (2021). Genre uptake as boundary-work: Reasoning about uptake in Wikipedia articles. *Journal of Technical Writing and Communication, 51*(2), 175-198.

Cronin, P., Ryan, F., & Coughlan, M. (2008). Undertaking a literature review: a step-by-step approach. *British Journal of Nursing; Br J Nurs, 17*(1), 38-43. https://doi.org/10.12968/bjon.2008.17.1.28059

Dang, Q., & Ignat, C.Measuring Quality of Collaboratively Edited Documents: the case of Wikipedia. Paper presented at the

Das, P., Guda, B. P. R., Seelaboyina, S. B., Sarkar, S., & Mukherjee, A. (2022). Quality change: norm or exception? Measurement, Analysis and Detection of Quality Change in Wikipedia. *Proceedings of the ACM on Human-Computer Interaction, 6*(CSCW1), 1-36.

Diggelmann, T., Boyd-Graber, J., Bulian, J., Ciaramita, M., & Leippold, M. (2020). Climate-fever: A dataset for verification of real-world climate claims. *arXiv Preprint arXiv:2012.00614,*

Diringer, E., & Perciasepe, B. (2020). The climate awakening of global capital. *Bulletin of the Atomic Scientists, 76*(5), 233-237.

Doyle, J., Farrell, N., & Goodman, M. K. (2017). Celebrities and climate change. *Oxford Research Encyclopedia of Climate Science* ()

Dunlap, R. E., & McCright, A. M. (2011). Organized climate change denial. *The Oxford Handbook of Climate Change and Society, 1*, 144-160.

Dunn, P. K., Marshman, M., & McDougall, R. (2018). Evaluating Wikipedia as a self-learning resource for statistics: You know they'll use it. *The American Statistician,*

Eichhorn, J., Molthof, L., & Nicke, S. (2020). From Climate Change Awareness to Climate Crisis Action. *Public Perceptions in Europe and the United States,*

Erxleben, F., Günther, M., Krötzsch, M., Mendez, J., & Vrandečić, D. (2014). Introducing Wikidata to the linked data web. Paper presented at the *International Semantic Web Conference,* 50-65.

Esteves Gonçalves da Costa, Bernardo, & Cukierman, H. L. (2019). How anthropogenic climate change prevailed: A case study of controversies around global warming on Portuguese Wikipedia. *New Media & Society, 21*(10), 2261-2282.

Esteves, B., & Cukierman, H. (2012). The climate change controversy through 15 articles of Portuguese Wikipedia. *Research Papers.Wikipedia Academy,*

Evans, R. (2021). *Wikipedia and Climate Misinformation.* Apple Podcasts. https://podcasts.apple.com/us/podcast/wikipedia-and-climate-misinformation/id1449762156?i=1000545733815

Fankhauser, S., Smith, S. M., Allen, M., Axelsson, K., Hale, T., Hepburn, C., Faulkner, R., Walling, S., & Pinchuk, M. (2012). Etiquette in wikipedia: Weening new editors into productive ones. Paper presented at the *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration,* 1-4.

Filatova, E. (2009). Directions for exploiting asymmetries in multilingual Wikipedia. Paper presented at the *Proceedings of the Third International Workshop on Cross Lingual Information Access: Addressing the Information Need of Multilingual Societies (CLIAWS3),* 30-37.

Ford, H., & Wajcman, J. (2017). 'Anyone can edit', not everyone does: Wikipedia's infrastructure and the gender gap. *Social Studies of Science, 47*(4), 511-527. https://doi.org/10.1177/0306312717692172

Forte, A., & Bruckman, A. (2008). Scaling consensus: Increasing decentralization in Wikipedia governance. Paper presented at the *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008),* 157.

Forte, A., Larco, V., & Bruckman, A. (2009). Decentralization in Wikipedia governance. *Journal of Management Information Systems, 26*(1), 49-72.

Gallagher, J. R., & Beveridge, A. (2021). Project-oriented web scraping in technical communication research. *Journal of Business and Technical Communication,* , 10506519211064619.

Häkkinen, K., & Akrami, N. (2014). Ideology and climate change denial. *Personality and Individual Differences, 70*, 62-65.

Hara, N., & Doney, J. (2015). Social construction of knowledge in Wikipedia. *First Monday,*

Henderson, R. M., Reinert, S. A., Dekhtyar, P., & Migdal, A. (2015). Climate change in 2018: Implications for business. *Risk, 1*

Hernández, D., Hogan, A., & Krötzsch, M. (2015). Reifying RDF: What works well with wikidata? *Ssws@ Iswc, 1457*, 32-47.

Higde, E., Oztekin, C., & Sahin, E. (2017). Turkish pre-service science teachers' awareness, beliefs, values, and behaviours pertinent to climate change. *International Research in Geographical and Environmental Education, 26*(3), 253-263.

Hinnosaar, M., Hinnosaar, T., Kummer, M., & Slivko, O. (2021). Wikipedia matters. *Journal of Economics & Management Strategy,*

Iordanou, K., & Constantinou, C. P. (2014). Developing pre-service teachers' evidence-based argumentation skills on socio-scientific issues. *Learning and Instruction, 34*, 42-57.

IPCC, C. W. T. (2007). *Climate change 2007: synthesis report*. IPCC Geneva, Switzerland.

Ivankova, N., & Wingo, N. (2018). Applying mixed methods in action research: Methodological potentials and advantages. *American Behavioral Scientist, 62*(7), 978-997.

Jamali, H. R., & Asadi, S. (2010). Google and the scholar: the role of Google in scientists' information‐seeking behaviour. *Online Information Review,*

Johansson, S., & Lindberg, Y. (2019). Wikipedia as a virtual learning site and a multilingual languaging site. *Virtual Sites as Learning Spaces,* , 181-203.

Johnson, R. (2011). Climate change education in K-12: Teacher preparation, understanding, needs and concerns. *Slingerlands, NY: National Earth Science Teachers Association,*

Kažys, J. (2016). The internet search results on the term 'climate change'by local Baltic Sea Region languages. Paper presented at the *Proceedings of Fifth International Conference on Climate Change Adaptation 2016 Challenges and Issues in Adaptation,* 23-30.

Kažys, J. (2018). CLIMATE CHANGE INFORMATION ON INTERNET BY DIFFERENT BALTIC SEA REGION LANGUAGES: RISKS OF DISINFORMATION & MISINTERPRETATION. *Journal of Security & Sustainability Issues, 7*(4)

Keegan, B., Gergle, D., & Contractor, N. (2011). Hot off the wiki: dynamics, practices, and structures in Wikipedia's coverage of the Tōhoku catastrophes. Paper presented at the *Proceedings of the 7th International Symposium on Wikis and Open Collaboration,* 105-113.

Keegan, B., Gergle, D., & Contractor, N. (2012). Staying in the loop: Structure and dynamics of Wikipedia's breaking news collaborations. Paper presented at the *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration,* 1-10.

Kenneally, E., & Dittrich, D. (2012). The Menlo Report: Ethical principles guiding information and communication technology research. *Available at SSRN 2445102,*

Khder, M. A. (2021). Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application. *International Journal of Advances in Soft Computing & its Applications, 13*(3)

Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher, 42*(8), 846-854.

Kim, S., Park, S., Hale, S. A., Kim, S., Byun, J., & Oh, A. H. (2016). Understanding editing behaviors in multilingual Wikipedia. *PloS One, 11*(5), e0155305.

Kincaid, D. W., Beck, W. S., Brandt, J. E., Brisbin, M. M., Farrell, K. J., Hondula, K. L., Larson, E. I., & Shogren, A. J. (2021). Wikipedia can help resolve information inequality in the aquatic sciences. *Limnology and Oceanography Letters, 6*(1), 18-23.

Kittur, A., & Kraut, R. E. (2008). Harnessing the wisdom of crowds in wikipedia: quality through coordination. Paper presented at the *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work,* 37-46.

Kivunja, C., & Kuyini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education, 6*(5), 26-41.

Koteyko, N., Nerlich, B., & Hellsten, I. (2015). Climate change communication and the internet: Challenges and opportunities for research. *Environmental Communication, 9*(2), 149-152.

Krange, O., Kaltenborn, B. P., & Hultman, M. (2019). Cool dudes in Norway: climate change denial among conservative Norwegian men. *Environmental Sociology, 5*(1), 1-11.

Kronthaler, F., & Zöllner, S. (2021). *Data Analysis with RStudio*. Springer.

Krotov, V., & Silva, L. (2018). Legality and ethics of web scraping.

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33: 159-174.

Lang, C. (2014). Do weather fluctuations cause people to seek information about climate change? *Climatic Change, 125*(3), 291-303.

Leas, E. C., Althouse, B. M., Dredze, M., Obradovich, N., Fowler, J. H., Noar, S. M., Allem, J., & Ayers, J. W. (2016). Big data sensors of organic advocacy: the case of Leonardo DiCaprio and climate change. *PloS One, 11*(8), e0159885.

Lee, T. M., Markowitz, E. M., Howe, P. D., Ko, C., & Leiserowitz, A. A. (2015). Predictors of public climate change awareness and risk perception around the world. *Nature Climate Change, 5*(11), 1014-1020.

Li, A., Yao, Z., Yang, D., Kulkarni, C., Farzan, R., & Kraut, R. E. (2020). Successful online socialization: lessons from the Wikipedia education program. *Proceedings of the ACM on Human-Computer Interaction, 4*(CSCW1), 1-24.

Liu, J., & Ram, S. (2018). Using big data and network analysis to understand Wikipedia article quality. *Data & Knowledge Engineering, 115*, 80-93.

Long, T., & Johnson, M. (2000). Rigour, reliability and validity in qualitative research. *Clinical Effectiveness in Nursing, 4*(1), 30-37.

Lörcher, I., & Taddicken, M. (2017). Discussing climate change online. Topics and perceptions in online climate change communication in different online public arenas. *Journal of Science Communication, 16*

Luís, S., Vauclair, C., & Lima, M. L. (2018). Raising awareness of climate change causes? Cross-national evidence for the normalization of societal risk perception of climate change. *Environmental Science & Policy, 80*, 74-81.

Luqman, Y. (2021). Millennials Information-Seeking Behavior About Climate Change.

McDowell, Z. J., & Vetter, M. A. (2020). It takes a village to combat a fake news army: Wikipedia's community and policies for information literacy. *Social Media Society, 6*(3), 2056305120937309.

McMahon, C., Johnson, I., & Hecht, B. (2017). The substantial interdependence of Wikipedia and Google: A case study on the relationship between peer production communities and information technologies. Paper presented at the *Eleventh International AAAI Conference on Web and Social Media,*

McNeal, P., Petcovic, H., & Reeves, P. (2017). What is motivating middle-school science teachers to teach climate change? *International Journal of Science Education, 39*(8), 1069-1088.

Menchen-Trevino, E., & Hargittai, E. (2011). YOUNG ADULTS'CREDIBILITY ASSESSMENT OF WIKIPEDIA. *Information, Communication & Society, 14*(1), 24-51.

Mesgari, M., Okoli, C., Mehdi, M., Nielsen, F. Å, & Lanamäki, A. (2015). "The sum of all human knowledge": A systematic review of scholarly research on the content of W ikipedia. *Journal of the Association for Information Science and Technology, 66*(2), 219-245.

Miquel-Ribé, M., & Laniado, D. (2020). The Wikipedia Diversity Observatory: A project to identify and bridge content gaps in Wikipedia. Paper presented at the *Proceedings of the 16th International Symposium on Open Collaboration,* 1-4.

Morgan, J. T., Bouterse, S., Walls, H., & Stierch, S. (2013). Tea and sympathy: crafting positive new user experiences on wikipedia. Paper presented at the *Proceedings of the 2013 Conference on Computer Supported Cooperative Work,* 839-848.

Moser, S. C. (2010). Communicating climate change: history, challenges, process and future directions. *Wiley Interdisciplinary Reviews: Climate Change, 1*(1), 31-53.

Moyer, D., Carson, S., Dye, T., Carson, R., & Goldbaum, D. (2015). Determining the influence of Reddit posts on Wikipedia pageviews. Paper presented at the *Proceedings of the International AAAI Conference on Web and Social Media, , 9*(5) 75-82.

Narayan, S., Orlowitz, J., Morgan, J., Hill, B. M., & Shaw, A. (2017). The wikipedia adventure: Field evaluation of an interactive tutorial for new users. Paper presented at the *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing,* 1785-1799.

Nordhaus, W. D. (2007). *The challenge of global warming: economic models and environmental policy*. Yale University New Haven.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1609406917733847.

Otero, P., López, I., Cilenis, S., & de Compostela, S. (2011). Measuring comparability of multilingual corpora extracted from wikipedia. *Iberian Cross-Language Natural Language Processings Tasks (ICL),* , 8.

Park, S. (2019). We love or hate when celebrities speak up about climate change: receptivity to celebrity involvement in environmental campaigns. *Journal of Contemporary Eastern Asia, 18*(1), 175-188.

Pillai, S. G., Soon, L., & Haw, S. (2019). Comparing DBpedia, Wikidata, and YAGO for web information retrieval. *Intelligent and Interactive Computing* (pp. 525-535). Springer.

Poudat, C., Vanni, L., & Grabar, N. (2016). How to explore conflicts in French Wikipedia talk pages? Paper presented at the *Statistics Analysis of Textual Data, , 2* 645-656.

Priedhorsky, R., Chen, J., Lam, S. (. K., Panciera, K., Terveen, L., & Riedl, J. (2007). Creating, destroying, and restoring value in Wikipedia. Paper presented at the *Proceedings of the 2007 International ACM Conference on Supporting Group Work,* 259-268.

Purcell, K., Rainie, L., & Brenner, J. (2012). Search engine use 2012.

Puschmann, C., & Ausserhofer, J. (2017). Social data APIs: Origin, types, issues.

Puttick, S., & Talks, I. (2021). Teachers' sources of information about climate change: A scoping review. *The Curriculum Journal,*

Rahi, S. (2017). Research design and methods: A systematic review of research paradigms, sampling issues and instruments development. *International Journal of Economics & Management Sciences, 6*(2), 1-5.

Redi, M., Fetahu, B., Morgan, J., & Taraborelli, D. (2019). Citation needed: A taxonomy and algorithmic assessment of Wikipedia's verifiability. Paper presented at the *The World Wide Web Conference,* 1567-1578.

Redi, M., Gerlach, M., Johnson, I., Morgan, J., & Zia, L. (2020). A taxonomy of knowledge gaps for wikimedia projects (second draft). *arXiv Preprint arXiv:2008.12314,*

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., & Waldman-Brown, A. (2022). Tackling climate change with machine learning. *ACM Computing Surveys (CSUR), 55*(2), 1-96.

Romm, N. R. (2018). Practicing Ethical Responsibility: Reconfiguring the Belmont Model. *Responsible Research Practice* (pp. 359-418). Springer.

Ross, A. D., Rouse, S. M., & Mobley, W. (2019). Polarization of climate change beliefs: the role of the millennial generation identity. *Social Science Quarterly, 100*(7), 2625-2640.

Rothshild, A., Lurie, E., & Mustafaraj, E. (2019). How the interplay of google and wikipedia affects perceptions of online news sources. Paper presented at the *Computation Journalism Symposium,*

Roy, D., Bhatia, S., & Jain, P. (2020). A topic-aligned multilingual corpus of Wikipedia articles for studying information asymmetry in low Resource languages. Paper presented at the *Proceedings of the 12th Language Resources and Evaluation Conference,* 2373-2380.

Roy, D., Bhatia, S., & Jain, P. (2022). Information asymmetry in Wikipedia across different languages: A statistical analysis. *Journal of the Association for Information Science and Technology, 73*(3), 347-361.

Saez-Trumper, D. (2019). Online disinformation and the role of wikipedia. *arXiv Preprint arXiv:1910.12596,*

Santana, A., & Wood, D. J. (2009). Transparency and social responsibility issues for Wikipedia. *Ethics and Information Technology, 11*(2), 133-144.

Saunders, M. N. K., Thornhill, A., & Lewis, P. (2019). Research Methods for Business Students (Eighth Edition ed.). Pearson Education Limited. 172, 182, 190, 193, 196, 208, 214, 215, 232, 266, 272, 435, 436, 646, 651,

Schäfer, M. S. (2012). Online communication on climate change and climate politics: a literature review. *Wiley Interdisciplinary Reviews: Climate Change, 3*(6), 527-543.

Schneider, J., Gelley, B. S., & Halfaker, A. (2014). Accept, decline, postpone: How newcomer productivity is reduced in English Wikipedia by pre-publication review. Paper presented at the *Proceedings of the International Symposium on Open Collaboration,* 1-10.

Scoones, I. (2007). Sustainability. *Development in Practice, 17*(4-5), 589-596.

Semenza, J. C., Hall, D. E., Wilson, D. J., Bontempo, B. D., Sailor, D. J., & George, L. A. (2008). Public perception of climate change: voluntary mitigation and barriers to behavior change. *American Journal of Preventive Medicine, 35*(5), 479-487.

Sepehri-Rad, H., & Barbosa, D. (2015). Identifying controversial Wikipedia articles using editor collaboration networks. *ACM Transactions on Intelligent Systems and Technology (TIST), 6*(1), 1-24.

Silva, M. (2021). *Climate change: Conspiracy theories found on foreign-language Wikipedia.* BBC. https://www.bbc.com/news/technology-59325128

Singer, P., Lemmerich, F., West, R., Zia, L., Wulczyn, E., Strohmaier, M., & Leskovec, J. (2017). Why we read Wikipedia. Paper presented at the *Proceedings of the 26th International Conference on World Wide Web,* 1591-1600.

Singrodia, V., Mitra, A., & Paul, S. (2019). A review on web scrapping and its applications. Paper presented at the *2019 International Conference on Computer Communication and Informatics (ICCCI),* 1-6.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333-339.

Theo van Veen. (2019). Wikidata. *Information Technology and Libraries, 38*(2), 72-81.

Thorne, S. (2000). Data analysis in qualitative research. *Evidence-Based Nursing, 3*(3), 68-70.

Thuiller, W. (2007). Climate change and the ecologist. *Nature, 448*(7153), 550-552.

Tol, R. S. (2009). The economic effects of climate change. *Journal of Economic Perspectives, 23*(2), 29-51.

Tran, M. (2009). Beyond bias: Wikipedia and the con-struction of knowledge. *Journal of Digital Research and Publishing, 3*(1), 48.

Ussiri, D. A., & Lal, R. (2017). *Carbon sequestration for climate change mitigation and adaptation*. Springer. 293

Varini, F. S., Boyd-Graber, J., Ciaramita, M., & Leippold, M. (2020). ClimaText: A dataset for climate change topic detection. *arXiv Preprint arXiv:2012.00483,*

Vetter, M. A. (2018). Teaching Wikipedia. *College English, 80*(5), 397-422.

Viégas, F. B., Wattenberg, M., Kriss, J., & Van Ham, F. (2007). Talk before you type: Coordination in Wikipedia. Paper presented at the *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07),* 78.

VijayaVenkataRaman, S., Iniyan, S., & Goic, R. (2012). A review of climate change, mitigation and adaptation. *Renewable and Sustainable Energy Reviews, 16*(1), 878-897.

Vincent, N., & Hecht, B. (2021). A deeper investigation of the importance of wikipedia links to search engine results. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW1), 1-15.

Vincent, N., Johnson, I., Sheehan, P., & Hecht, B. (2019). Measuring the importance of user-generated content to search engines. Paper presented at the *Proceedings of the International AAAI Conference on Web and Social Media, , 13* 505-516.

Voigt, P., & Von dem Bussche, A. (2017). The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing, 10*(3152676), 10.5555.

Vrandecic, D. (2018). Capturing meaning: Toward an abstract Wikipedia.

Vrandečić, D., & Krötzsch, M. (2014). Wikidata: a free collaborative knowledgebase. *Communications of the ACM, 57*(10), 78-85.

Waller, V. (2011). Not just information: Who searches for what on the search engine Google? *Journal of the American Society for Information Science and Technology, 62*(4), 761-775.

Warncke-Wang, M., Ranjan, V., Terveen, L., & Hecht, B. (2015). Misalignment between supply and demand of quality content in peer production communities. Paper presented at the *Proceedings of the International AAAI Conference on Web and Social Media, , 9*(1) 493-502.

Warrens, M. J. (2015). Five ways to look at Cohen's kappa. *Journal of Psychology & Psychotherapy, 5*(4), 1.

Whitmarsh, L., & Capstick, S. (2018). Perceptions of climate change. *Psychology and climate change* (pp. 13-33). Elsevier.

Wikimedia Foundation. (2021). *Terms of Use.* Terms of Use - Wikimedia Foundation Governance Wiki

Wikipedia. (2022). *2022 Russian invasion of Ukraine: Revision history.* Wikipedia.https://en.wikipedia.org/w/index.php?title=2022_Russian_invasion_of_Ukraine&offset=20220224040035%7C1073704451&limit=50&action=history

Wilkinson, D. M., & Huberman, B. A. (2007). Cooperation and quality in wikipedia. Paper presented at the *Proceedings of the 2007 International Symposium on Wikis,* 157-164.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. ACM. https://doi.org/10.1145/2601248.2601268

Wulczyn, E., West, R., Zia, L., & Leskovec, J. (2016). Growing wikipedia across languages via recommendation. Paper presented at the *Proceedings of the 25th International Conference on World Wide Web,* 975-985.

Xu, H., Zhang, N., & Zhou, L. (2020). Validity concerns in research using organic data. *Journal of Management, 46*(7), 1257-1274.

Yoshida, M., Arase, Y., Tsunoda, T., & Yamamoto, M. (2015). Wikipedia page view reflects web search trend. Paper presented at the *Proceedings of the ACM Web Science Conference,* 1-2.

Appendices

Appendix A <Intitial Literature List >

Appendix B < Transcript of the Interview 1>

Appendix C <Transcript of the Interview 2 >

Appendix D <Transcript of the Interview 3 >

Appendix E< Thematic Analysis Interview 1>

Appendix F <Thematic Analysis Interview 2 >

Appendix G <Thematic Analysis Interview 3>

Appendix H < Final Thematic Analysis>

Appendix I < Thematic Anlysis Themes and Sub-themes and their repetitions >

Appendix A < Intitial Literature List >

Literature Search

Appendix B <Interview Guide >

Interview Guide

Appendix C < Transcript of the Interview 1 >

Appendix C - Transcript Interview 1

Appendix D < Transcript of the Interview 2 >

Appendix D - Transcript Interview 2

Appendix E < Transcript of the Interview 3 >

Appendix E - Transcript Interview 3

Appendix F< Thematic Analysis Interview 1>

Appendix F Interview 1 - Thematic Analysis

Appendix G<Thematic Analysis Interview 2 >

Appendix G Interview 2 - Thematic Analysis

Appendix H <Thematic Analysis Interview 3>

Appendix H Interview 3 - Thematic Analysis

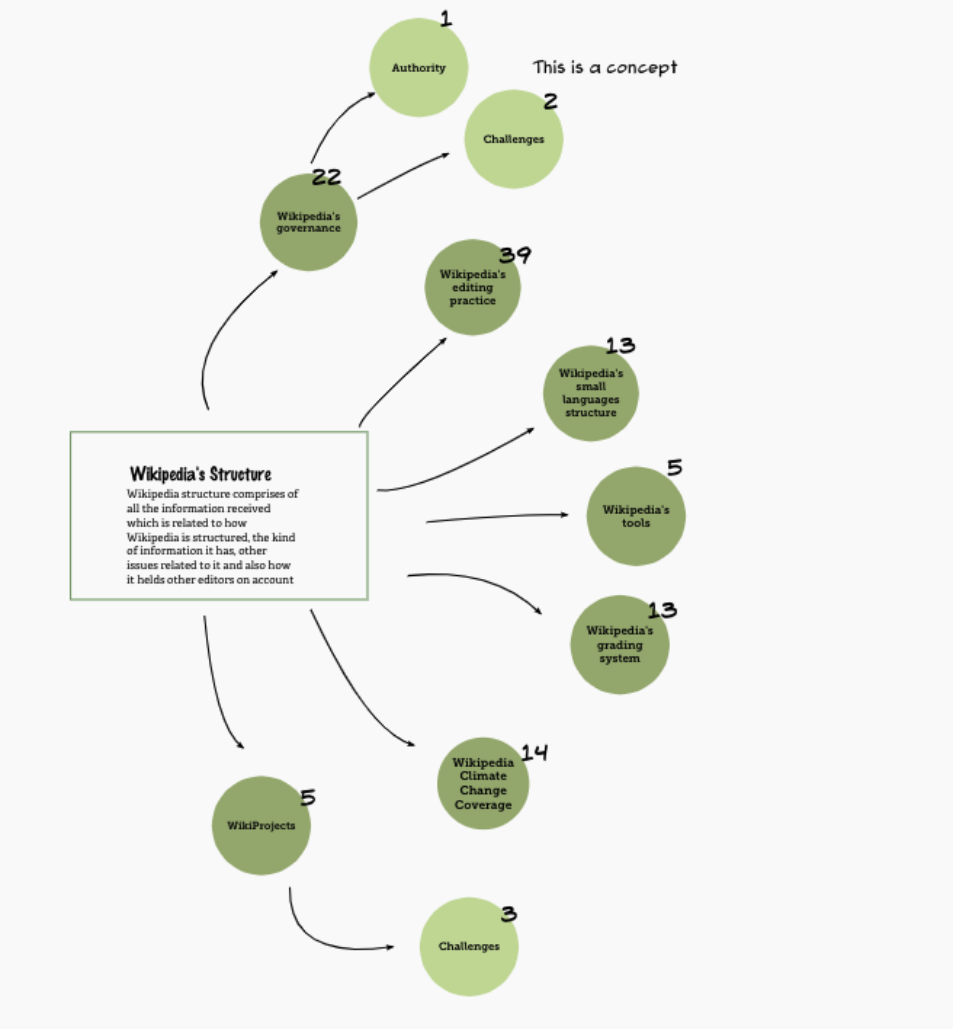
Appendix I< Final Thematic Analysis >

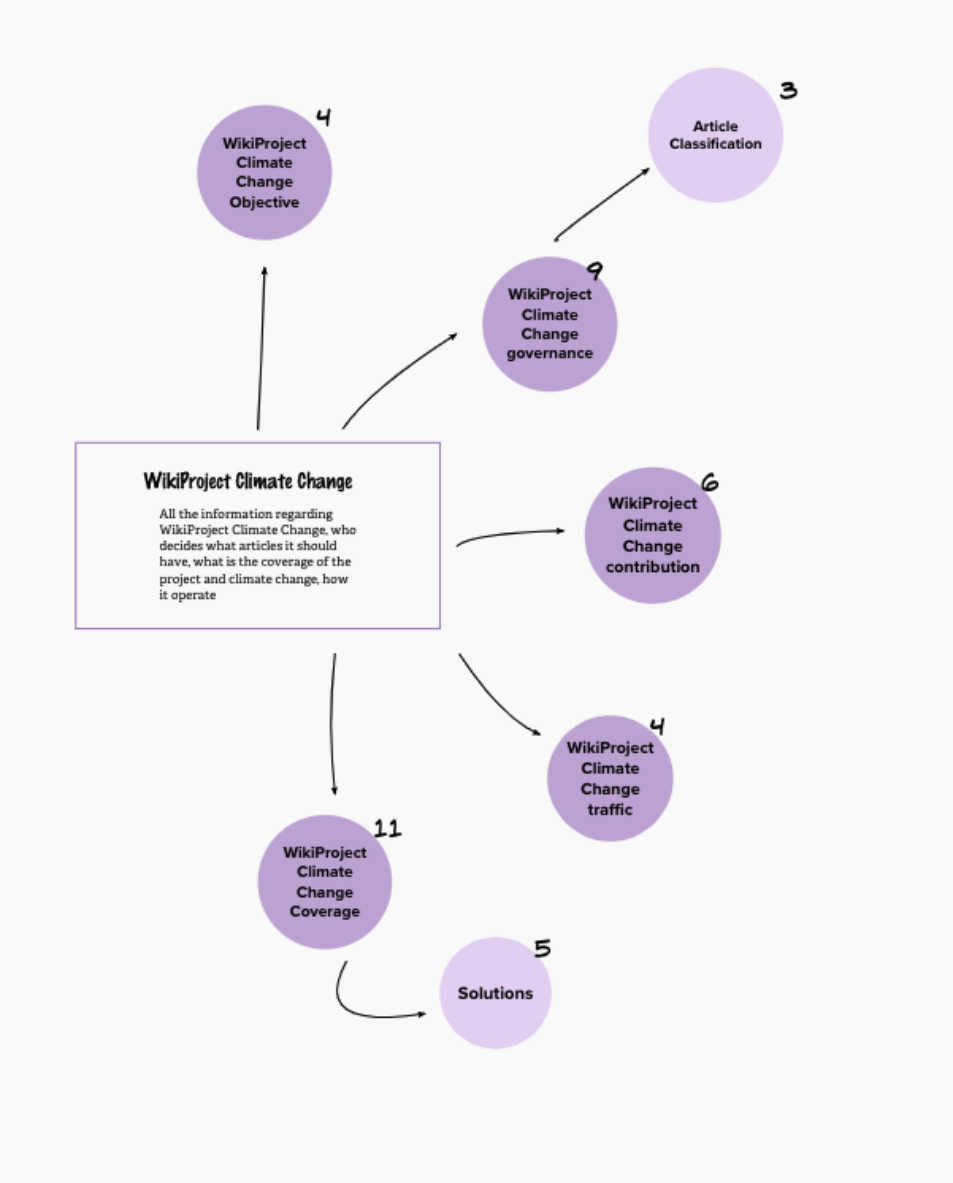
Final Thematic Analysis

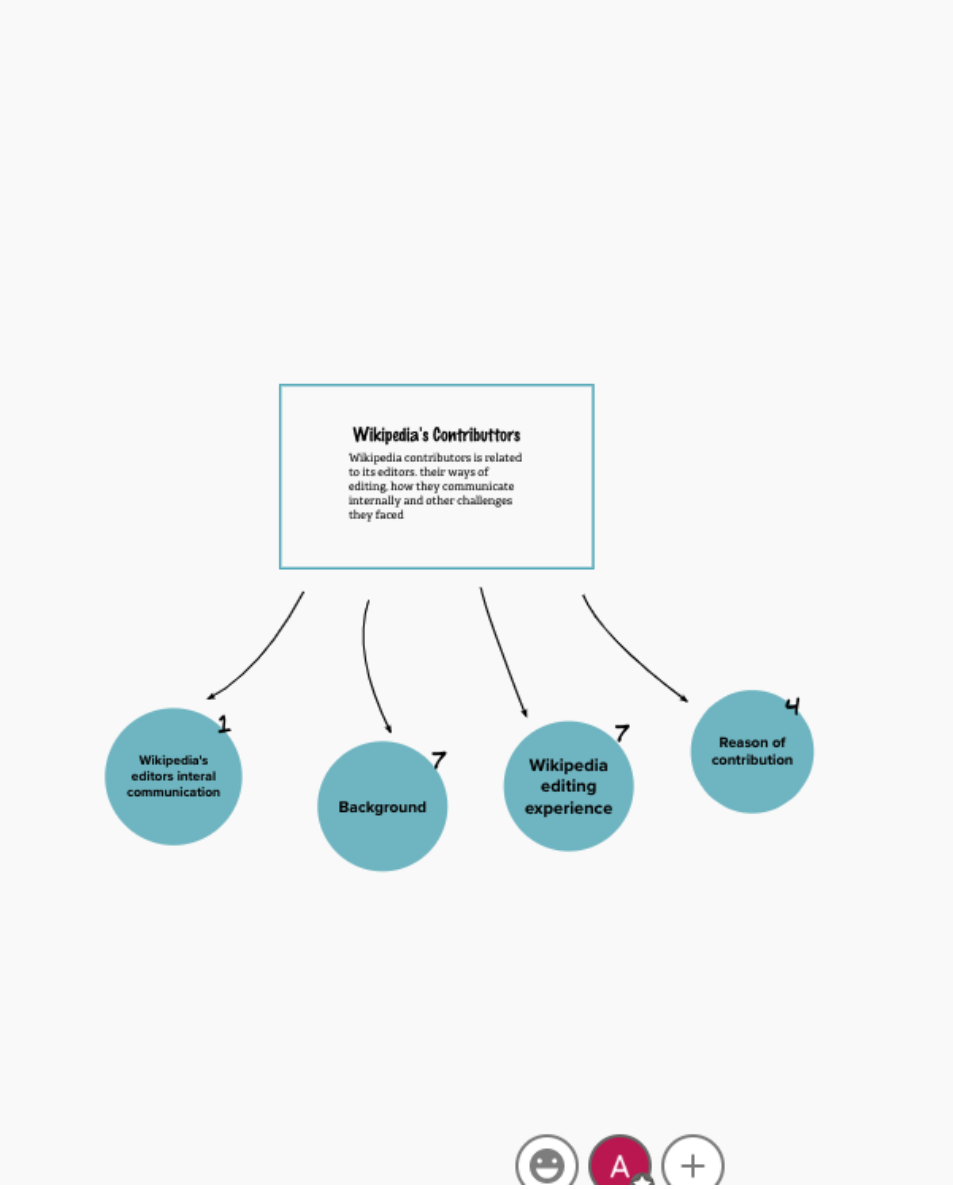
Appendix J<The Cohen’s Kappa Calculation>

Cohen's Kappa.xlsx

Appendix K <Thematic Analysis, Themes, Sub-themes, and their repetitions>




WikiProject Climate Change
All the information regarding WikiProject Climate Change, who decides what articles it should have, what is the coverage of the project and climate change, how it operate
4
WikiProject Climate Change Objective
3
Article Classification
9
WikiProject Climate Change governance
6
WikiProject Climate Change contribution
4
WikiProject Climate Change traffic
11
WikiProject Climate Change Coverage
5
Solutions


Wikipedia's Contributtors
Wikipedia contributors is related to its editors, their ways of editing, how they communicate internally and other challenges they faced
1
Wikipedia's editors interal communication
7
Background
7
Wikipedia editing experience
4
Reason of contribution