

UNIVERSITY OF WEST BOHEMIA

FACULTY OF ECONOMICS

Diploma thesis

**Projects related to the SDGs and the response generated
on social networks: analysis of the World Economic
Forum profile on Twitter over the selected time period**

**Projekty související s SDGs a jejich ohlas na sociálních
sítích: analýza profilu World Economic Forum na
Twitteru za vybrané období**

David Klika

Pilsen 2022

Prohlašuji, že jsem diplomovou práci na téma

„Projects related to the SDGs and the response generated on social networks: analysis of the World Economic Forum profile on Twitter over the selected time period“

vypracoval samostatně pod odborným dohledem vedoucího diplomové práce za použití pramenů uvedených v příložené bibliografii.

Plzeň dne 23. 4. 2022

v. r. David Klika

I would first like to thank my supervisor, Dr. Jiří Preis, whose expertise and feedback were invaluable in formulating the research questions and methodology.

I would also like to thank Dr. David Vogt, who provided me with crucial advice about the GIS software used in this thesis.

In addition, I would like to thank my mom and friends, but especially my friend Jessa. You were always there for me and believed in me. You also made sure I kept on track.

Last but not least, I would like to thank the YouTube channel Lofi Girl, which music kept me company through the long winter nights.

Table of Content

Introduction	7
1 Sustainability and the online environment	9
1.1 Introduction to sustainability	9
1.1.1 The geography of economy	9
1.1.2 Globalization, innovation and technological progress	11
1.1.3 Sustainable development, history and definitions	14
1.2 United Nations and sustainable development.....	16
1.2.1 United Nations	16
1.2.2 Millennium Development Goals.....	17
1.2.3 Sustainable Development Goals	19
1.3 Online environment, the Internet and Twitter	21
1.3.1 The Internet and its demography	21
1.3.2 Birth of the social media.....	23
1.3.3 Twitter, the microblogging social platform.....	25
2 Review of available literature	28
2.1 Sustainability	28
2.2 Sustainable Development Goals and the World Economic Forum.....	29
2.3 Social media and sustainability.....	30
3 Methodology and goals	31
3.1 Data formats and technology	31
3.1.1 JSON and JSONL	31
3.1.2 CSV	32
3.1.3 Snsrape.....	33
3.1.4 Java.....	34

3.1.5	Google Sheets	35
3.1.6	ESRI ArcMap	35
3.2	Goals.....	36
4	The World Economic Forum and its online presence	38
4.1	History and main ideas	38
4.2	Online presence.....	40
4.3	Twitter account	41
5	Data acquisition, description and basic analysis	44
5.1	Data acquisition.....	44
5.1.1	Twitter Advanced search.....	45
5.1.2	Data scraping	47
5.2	Descriptive statistics of WEF Twitter profile	49
5.3	Analysis of post frequency	50
6	Content analysis and results.....	53
6.1	Geographical content.....	53
6.1.1	Countries and territories	54
6.1.2	Continents.....	57
6.1.3	Other possible classifications	60
6.2	Topical and context approach.....	63
6.2.1	Popular hashtags	63
6.2.2	Popular user mentions	65
6.2.3	Other interesting hashtags or mentions	70
6.3	Most popular tweets	71
6.3.1	Top 10 liked tweets.....	71
6.3.2	Top 10 retweeted tweets.....	72

6.3.3	Top 10 replied to tweets	74
6.3.4	Top 10 quote retweeted tweets	77
7	Conclusions, findings and further work	81
7.1	Geographical content.....	81
7.2	Popular and important topics	82
7.3	Most popular tweets	83
7.4	Overlapping meanings of keywords.....	84
7.5	Repetitiveness of tweets	84
7.6	Suggestions for further work	85
	Conclusion	86
	List of the Bibliographic References.....	88
	List of Tables	94
	List of Figures.....	95
	List of Abbreviations and Symbols.....	96
	List of Appendices	97
	Appendix A: Full table with country data	
	Abstract	
	Abstrakt	

Introduction

Sustainability and sustainable development are the topics standing at the forefront of political debates, corporate and business efforts and which found their way into the public discourse around the world in the past years. In 2015 United Nations General Assembly announced their mission for the better world by 2030. By signing the resolution commonly known as the Agenda 2030 they presented 17 Sustainable Development Goals (SDGs), along with 169 targets and further indicators, following in the footsteps of their previous Millennium Development Goals. The 17 SDGs have quickly become popular and used metric for sustainability by the multitude of actors, from transnational organization to small business. Similarly, the Internet and the rapid spread of modern technologies across the globe are bringing more and more people into the online environment. In its current form the World Wide Web, Internet's most used technology, provides space for a multitude of platforms where individuals, organizations and governments can share their thoughts, interact with each other and discuss the issues of the contemporary world, including sustainability.

The main goal of this thesis is to examine one such example, the World Economic Forum and the way they present themselves on the Internet, particularly on the popular social media platform Twitter. Our aim is to analyze the content posted onto their Twitter account over the selected period of 5 years, from 2016 to 2020, and present relevant results focusing on the topics of geographical content, topical content and the most popular content. We put special emphasis on the topic of sustainability and sustainable development, tying the posted content to the SDGs as a possible metric for sustainability in public discourse.

In the first chapter of this thesis we lay the necessary theoretical foundation for sustainability, SDGs and the development of the online environment and the social media. In the second chapter we examine and explore the contemporary literature on sustainability, SDGs and sustainability in the context of social media. Third chapter explained the used methodology and concretize the goals of this thesis. Fourth chapter describes the context of the World Economic Forum and its presence online. Fifth focuses on the acquisition of the relevant data, describes the techniques to more detail and provides the initial analysis. Sixth chapter analyses the acquired dataset deeper,

with the emphasis on geographical content, topical content and the most popular content and presents relevant results. The seventh and last chapter summarizes previous chapters, discusses the results and provides incentives for future and further work.

1 Sustainability and the online environment

When one is to talk about sustainability and sustainable development, there are certain concepts and terms, which tie closely to the problematic, they inevitably encounter, and which need to be properly explained. As with many other concepts within the field of social studies these are at times used rather loosely and their meaning change depending on the context; be that author, publication, or concrete field.

In the forefront of the global sustainability efforts stands the United Nation's (UN) mission which introduced the 17 Sustainable Development Goals (SDGs).

Further, as this thesis focuses on sustainability in the online environment, in particular on Twitter, one of the leading social networking platforms, it is important to explain the basic principles of online communication and functionality available through the platform.

Thus, this chapter aims to clarify important concepts related to sustainability within the context of sustainable development. First, it takes a step back to describe the underlying concepts of economy and its organization within space and provides important historical background to sustainability. Further it describes the role of the UN as a major actor for sustainability along with their projects of Millennium Development Goals (MDGs) and SDGs. Finally, it describes the broader context of the online environment – the Internet and its demography, what allowed for the origin of the social media, and the functionality of Twitter.

1.1 Introduction to sustainability

1.1.1 The geography of economy

If we are to talk about sustainability, it is important to pivot such concepts within a broader framework of the *economy* and its development throughout history. It is a science related to *how* people *produce, consume, trade, and distribute goods and services* among each other. However, plainly asking “how” would be an oversimplification. All the actions within economy are closely connected to the *geography* of the area they take place at. To understand the *spatiality of economy* more

closely, it is important to distinguish between its individual aspects of *location*, *territory*, *place*, and *scale*.

Location involves the position of people and objects relative to each. To get between locations one must surmount the *distance*, which requires time and money expenditures. These costs in turn determine the layout of the economic landscape. Further, there is an important distinction between absolute and relative distance. The *absolute distance* relates strictly to how far two points are spatially within a certain coordinate system (ex. latitude and longitude). This approach helps with the basic understanding. However, sometimes it is important to consider the underlying infrastructure of certain regions. That, taking transport cost and travel time into consideration, is the *relative distance*. The idea of location and distance is crucial for the works of early economic geographers and their models. The most notable examples are Johann-Heinrich von Thünen and his model of *isolated state* and Walter Christaller and his *central place theory* (Coe et al., 2013).

Second, *territory* relates to the factual power someone is able to exercise over a certain part of land. The primary example of territorial power is governments, who can steer the economic activities in their territories and those flowing across their border, from their territory to another (Coe et al., 2013).

Third, *places* are geographical areas with cultural and political significance, which provides them with certain uniqueness when compared to other places. Such characteristics pose importance to those who interact with them, even from the economic point of view as they provided the opportunity for unique economic outcome. This uniqueness of outcomes in turn determines whether and what economic activities will be held in such places (Coe et al., 2013).

Lastly, the *scale* of a region is an important concept for economic activities. We can address certain issues from the point of view of a particular household, city, county, country, macroregion, or even the whole world. These were some of the examples of geographical scale (Coe et al., 2013).

Further in this thesis we focus predominantly on the concept of territories, which are most often presented by individual countries. However, places and scale is also considered. The places important for the World Economic Forum (WEF) in particular

should be examined. The scale of a region is considered implicitly, when we work with regions of various sizes. However, as described further in the work, the problematic of SDGs is aimed at global scale, but its indicators and targets can be addressed on country or local level as well.

1.1.2 Globalization, innovation and technological progress

Now that we understand how economy and geography are connected, we may take a look at the development of factors influencing economic geography on a larger scale throughout history.

For the long time the idea of economy and the economical thinking was a highly localized idea. It included people's households and an access to local market. There people exchanged their products with relatively low over-production for money or goods. The main goal of production was to first cover my own needs and then the needs of my local neighbors. The division of labor was present at the time, but was not as specialized (Coe et al., 2013).

The main force driving the division of labor further was the Industrial Revolution (IR), a major leap in technological progress. So far three such IR occurred, with the fourth being under way.

The First IR in the late 18th century was tightly connected to mechanization and introduction of steam power. It had pushed the handicraft economy aside and agricultural production and manufacturing emerged on a much larger scale than before. The manufacturing was held in large factories and led to further and more fragmented division of labor (Xu et al., 2018). This rapid change in the speed and scale of production has led to surplus, which in turn provided the opportunity for large scale trading (Coe et al., 2013). The Second IR in the late 19th century was signified by introduction of electricity and electrification, which in turn led to massive assembly lines and mass production. Then, the Third IR in the 1970s signaled progress in automation and computing technologies, along with the introduction of micro-electronics. While the first three IR had followed almost a century after each other, the Fourth IR, coined by Klaus Schwab, is a phenomenon of the 2010s and further. It is IR underway and it is shaping the world around us. It is based on the progress within the informatics and communication technologies (Xu et al., 2018). It allows for the

connection of people and systems across the globe through technologies such as the Internet and the World Wide Web. The progress in these technologies gives us the ability to acquire and process large amount of data for analysis and precise predictions based on them. It is a complex process that is reshaping all facets of our world, from economy, through government, to the very core of our society and identity (Springer & Schnelzer, 2019). The 4th IR comes with many other modern “4.0” aspects – Industry 4.0, World 4.0 and the adjective of *smart* and *intelligent*. We have smart materials, systems and technologies. Overall, it’s the combination of sensors and powerful computers to *smartly and intelligently* interact with its surrounding to help humans. From army and defense, to smartphones, smartwatches, smart lights and smart fridges (Goddard et al. 1997). With those come the ideas for smart cities – vast utilization of smart technologies, with smart lights in buildings and smart screens showing information about bus and trains arrivals and departures.

One factor is the key feature of all of the IRs. It is technological progress achieved through innovation. However, those innovations are not spread equally throughout the world. As time and technology progressed, it had brought an ever increasing division of labor. First it was on the level of a factory, where each worker focused on one highly optimized task. But as the progress in transporting technologies lowered the relative distances and costs of travel, it propelled the division of labor to the global scale (Dicken, 2015). We live in a world, where the raw material may come from Southern Africa, is transported to Southeast Asia for initial processing, then to Eastern Europe for coloring, and finally is sold to the customers in North America while profit is still made along the way. With global economy this connected and tangled, many questions about its real efficiency, fairness and transparency arise and some are definitely addressed by the UN’s SDGs and other projects and metrics of sustainability – certification, corporate social responsibility and so on (Coe et al., 2013). Further, the ability to share crucial information about the state of a region, company or individual within seconds provides a powerful tool for platforms for political agendas, both from the top (governments, companies, producers), and the bottom (citizens, consumers) (Dicken, 2015).

This interconnectedness of society and economy across the globe leads us to another important topic – *globalization* and its impact on the world around us.

But, what even is globalization? Historically several authors and documents noticed the growing spatiality and interconnectedness of the economic activities. The word itself has been sparsely used in the English language since the early 20th century. However, it became commonly used in the last decades as a notion of economy “to become global”. This includes two important notions – *spatial spread of economic activities* and *growing functional integration of economic activities*. With time our ability as individuals to understand of the global chains of production and the flow of materials and goods became more complicated, as the relationships within and between companies became more interconnected and entangled (Dicken, 2015).

At the beginning of the new century globalization felt as something natural and inevitable. However, the global economic crises which had begun in 2008 struck a huge blow into this way of seeing globalization. As such we can identify multiple points of view regarding globalization.

First such approach, stemming from the previous ideas, is the hyper-globalists. Those claim that the world of today is borderless, nation-states and nationalities no longer have meaning, cultural and societal differences are diminished in face of the globalized economy. The products of this economy are homogeneous and uniform, bearing no marks of their cultural and geographical origins. In the words of American writer and political commentator Thomas Friedman: “The World is flat.” Relative distances across the globe are much shorter. We can get ourselves, material and goods from one side of the world to the other within days or hours. Further, information can be spread within seconds. By one press of a button a virtual meeting can be held in real-time between people from almost anywhere on this planet (Dicken, 2015).

However, the pure hyper-globalist approach has not yet been achieved and the question is begged to be asked – will it be ever possible? Other academics, thinkers and politicians would disagree. For this purpose, it is important to pin down the underlying political ideas. There is a significant dichotomy in the perception of globalization from the politicians on the right and left side of the political spectrum.

While the political right – neo-liberals, materialized in the hyper-globalist, advocate for an open and free-market economy and argue that globalization is a solution to the issues

of today's world. In their eyes we should let it develop further naturally without restrictions and in turn economic growth occurs and the quality of life improves.

The left-leaning thinkers – socialists and social democrats, the anti-globalists, see globalization as the root of the global issues. They argue for market regulations, otherwise the inequality among people will grow. The spatially spread and hard-to-look-into operations allows for the drainage of resources from one region to another, further widening the economic gap between the regions – the rich get richer, the poor stay poor. These anti-globalists stand opposite to the hyper-globalists and see the solution to this growing disproportion in abolishing the global and in returning to the local (Dicken, 2015).

1.1.3 Sustainable development, history and definitions

When the topic of sustainability is brought up, one may have an intuitive understanding of what it encompasses. However, throughout the years the term itself, as well as the accompanying term of sustainable development, evolved in its meaning targeting various aspects of the world around us.

The first proper definition of “sustainable development” dates way back to 1713 when Hans Carl von Carlowitz published his book on forest sciences called *Sylvicultura Oeconomica*. In this publication the author speaks about a concept called “*Nachhaltigkeit*”, which can be translated to English as “sustainability”. In this book Carlowitz describes the threatening shortage of available timber. Thus, Carlowitz proposes that local forests and their timber should be “*used with caution in a way that there is balance between timber growth and lumbering*” (Carlowitz, 1713). He further adds, that ““*[f]or this reason, we should organize our economy in a way that we won't suffer scarcity [of timber], and where it is lumbered we should strive for young growth at its place*” (Carlowitz, 1713). Within the context of this publication Carlowitz adds an important implicit layer to sustainability - growing timber is a long-term, multi-generational process (Carlowitz, 1713).

Based on this initial definition of sustainable development by Carlowitz is the definition from the 1987's World Commission of Environment and Development (known as the *Brundtland Commission*): “*Sustainable development meets the needs of the present generation without compromising the ability of future generations to meet their needs.*”

This definition of sustainability is often used verbatim or with slight variance in public discourse and media and by politicians (Keiner, 2005).

However, over the years the aforementioned definition found itself criticized for the lack of clarity and lack of concrete goals and topics to follow. Thus, the Swiss “Monitoring of Sustainable Development Project” MONET redefined the previous Brundland’s definition using concrete key factors. Their proposed definition:

“Sustainable development means ensuring dignified living conditions with regard to human rights by creating and maintaining the widest possible range of options for freely defining life plans. The principle of fairness among and between present and future generations should be taken into account in the use of environmental, economic and social resources.

Putting these needs into practice entails comprehensive protection of biodiversity in terms of ecosystem, species and genetic diversity, all of which are the vital foundations of life” (MONET in Keiner, 2005).

With this definition of sustainable development another important concept has emerged. The *environmental, economic and social resources* are defined, which are today often considered as the *three pillars of sustainable development*, individual aspects to sustainability which form the stable base for its success (Keiner, 2005).

Those form the three pillar model of sustainable development. However, Keiner (2005) further mentions other possible approaches to sustainable development: Capital Stocks, Prism models and the Egg model.

However, all these definitions have certain common features. First is the longevity of the problematic. It does not focus on the issues of one generation, but rather on the issues which are intergenerational, with the idea of “passing on” this world to the next generation. That is to be done in at least as good of a state as it had been for the generation before. Second, sustainable development is anchored by (at least) three interconnected aspects – environmental, economic, and societal. Issues related to these aspects then need to be carefully balanced by the decision-makers on all levels of governance - international, national, regional, and local. Third, the key actors for this problematic are recognized - the United Nations, governments, private sector, and non-governmental organizations (Mensah, 2019).

1.2 United Nations and sustainable development

When one is to talk about, or read into the topic of sustainable development it is easy to encounter three words - Sustainable Development Goals. This chapter focuses on describing the SDGs, their history in preceding projects of MDGs and the overall history of the UN and their sustainability efforts.

1.2.1 United Nations

The history of the United Nations dates back to the year 1945. It was then, while World War II was still ongoing, when the members of 50 world's governments met for a conference which resulted in the first draft of the UN Charter. This Charter, which served as the founding document for the organization, was signed on 26 June 1945 and came into force on 24 October 1945 and with it the UN was established and began its operations. The organization strived to be successful and prevalent where its predecessor, League of Nations, had failed. The then-goal of the UN was to promote and maintain international peace and security, as well as the upholding of human rights throughout the world.

However, as years passed the UN grew and transformed. The 51 founding states expanded into, as of today, 193 member states, with additional 2 non-member observer states. Similarly, its goals and focus have moved on to encompass a broader variety of topics, including humanitarian aid, international law, global issues and sustainable development and climate action. In and of itself, the UN System consists of six principal organs, several Specialized Agencies and related organizations. To mention a few - *Food and Agriculture Organization (FAO)*, *International Monetary Fund (IMF)*, *World Bank Group (WBG)*, *World Health Organization (WHO)*, and *United Nations Educational, Scientific and Cultural Organization (UNESCO)*.

It was thus perhaps a natural crystallization of the contemporary topics and discussions of the 1990s and the new millennium, when the UN came with the campaign for their initial idea of MDGs, the first iteration of comprehensive ideas connecting the economic, social and environmental views about the world, defining the most grievous of the world's issues (United Nations [UN], n.d.a).

1.2.2 Millennium Development Goals

The year 2000 posed both an important and symbolic moment in time. It was, from the etymological point of view quite literal, a milestone for the society of the world. At the time the UN's General Assembly held an important meeting, which as a result brought two important ideas for the future of economic development and sustainability.

The first of which was the adoption of the United Nations Millennium Declaration on the September 8th of 2000. There the 189 member states expressed their vision and outlook into the 21st century, as well as describing certain core problems of the time. In this document core values and principles for the following century are proposed - freedom, equality, solidarity, tolerance, respect for nature, and shared responsibility. Building on those, as a general baseline are further chapters focusing on general statements about world's issues. Among the topics we see - *peace, security and disarmament; development and poverty eradication; protecting our common environment; human rights, democracy and good governance; protecting the vulnerable; meeting special needs for Africa; and strengthening the United Nations*. Each of the aforementioned contains a brief introduction to the problematic, followed by a list of general resolutions about each topic for the years to come (United Nations General Assembly [UNGA], 2000).

The second was the proposition of the MDGs, which were adopted by not only the member states, but also by more than 20 international organizations, such as the WBG, or WHO. These goals were closely related to the general ideas proposed in the aforementioned Millennium Declaration. Focusing on human capital, infrastructure and human rights they established 8 goals with 21 targets measured by 60 indicators aimed to provide a generic, but measurable framework for tracking the progress towards each of the goals on the national and global level between the years 1990 and 2015, thus over the time period of 25 years (Liverman, 2018).

The goals themselves aimed to:

1. Eradicate extreme poverty and hunger,
2. Achieve universal primary education,
3. Promote gender equality and empower women,
4. Reduce child mortality rates,
5. Improve maternal health,

6. Combat HIV/AIDS, malaria, and other diseases,
7. Ensure environmental sustainability,
8. Develop a global partnership for development (United Nations Development Programme [UNDP], 2010).

As mentioned before, each of these goals has more specific targets, which are then measured by, predominantly health and economic indicators. For example Goal 1, Eradicate extreme poverty and hunger, has in total 3 targets, one of which is *Target 1A: Halve, between 1990 and 2015, the proportion of people living on less than \$1.25 a day*. This target is then measured by two indicators: *Poverty gap ratio (incidence x depth of poverty)*, and *Share of poorest quintile in national consumption*. All of the 8 goals are broken down in a similar way (UNDP, 2010).

Further, there are other important influencing factors about the MDGs.

First is the *mutual interconnectedness of goals* among themselves. The impact of individual decisions on the MDGs should be assessed in regard to all of them or at least to the relevant group. The over-focusing on one particular MDG may lead to deterioration in the progress of the others (Liverman, 2018). The same idea is then true for the SDGs as those should be also assessed together or in consideration with the rest, as some goals supplement each other well, while others work more as a counterforce (Nilsson et al., 2016).

Second is the *scalability* of the MDGs, which tie closely into the problematic of overall scale in economic geography. Proposed goals offer a general framework within which each actor operates with their individual focuses and needs. Thus, it is important to scale down the MGDs accordingly to national, regional, and local levels. This should lead to the creation of further documents and partial targets by each interested actor from nation-wide plans, to regional and local plans well-fitted for that particular area, its demographic, infrastructure and needs (Liverman, 2018).

Lastly is the *time frame* for the MDGs, the 25 years between 1990 and 2015. While the discussions about certain topics were ongoing through the 1990s, it is important to note that the goals themselves were adopted in 2000. Hence, the MGDs were adopted ten years into the project's proposed time frame. Some critics argue that it was a cynical and calculated move to weaken the MDGs, as some progress towards the goals (especially in rapidly growing China) was already made. The counter argument here is

the overall lack of reliable data for the year 2000 as a baseline by the time of the Millennium Summit, as well as the ongoing global trends stemming from the 1990s and the aforementioned topical conferences which often shared the baseline of 1990 (Liverman, 2018).

1.2.3 Sustainable Development Goals

While 2015 marked the end of the MDGs, where the UN drew their conclusions in their report (UN, 2015a) to at least partial success for most of its goals and targets, they also recognized the growing scope of sustainable development problematic. Thus, beginning in 2015 they adopted new mission and a new set of goals, both broader and deeper in its topics. The document commonly known as *2030 Agenda* was adopted during the 70th session of the UNGA and introduced 17 new goals for sustainable development with the target year of 2030 – the *Sustainable Development Goals* (UN, 2015b).

The 17 SDGs are the following:

- Goal 1: End poverty in all its forms everywhere
- Goal 2: End hunger, achieve food security and improved nutrition and promote sustainable agriculture
- Goal 3: Ensure healthy lives and promote well-being for all at all ages
- Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
- Goal 5: Achieve gender equality and empower all women and girls
- Goal 6: Ensure availability and sustainable management of water and sanitation for all
- Goal 7: Ensure access to affordable, reliable, sustainable and modern energy for all
- Goal 8: Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all
- Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation
- Goal 10: Reduce inequality within and among countries
- Goal 11: Make cities and human settlements inclusive, safe, resilient and sustainable
- Goal 12: Ensure sustainable consumption and production patterns
- Goal 13: Take urgent action to combat climate change and its impacts
- Goal 14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development

- Goal 15: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss
- Goal 16: Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels
- Goal 17: Strengthen the means of implementation and revitalize the global partnership for sustainable development (UN, 2015b).

Along with the main group of 17 goals a list of 169 more concrete targets was included. Further, each target includes measurable indicators, with 1 to 4 indicators per target. These global indicators were developed by the Inter-Agency and Expert Group on SDG Indicators and later adopted by the UN's Statistical Commission in 2017. In March 2020 some of the indicators were updated and refined (UN, n.d.b).

Usually, when the targets or indicators are mentioned they are identified with their corresponding goal through multi-level labels. The same labeling system is used throughout this thesis as well. Individual targets are either numbered, or marked by a letter depending on the target subtopic. Indicators are always numbered.

As an example of a target and its indicators the Goal 1 has a total of 7 targets, 5 of which are numbered, 2 of which are marked by a letter. The target 1.2 reads: *“By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions”*. To measure its progress and success two indicators are used: *“1.2.1 Proportion of population living below the national poverty line, by sex and age”* and *“1.2.2 Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions”* (United Nations Statistics Division [UNSD], n.d.).

It is important to note that many of the data collected is done on a national level and through UN affiliated agencies, which is to provide a unifying factor for the data structure. Individual goals themselves are often highly interconnected which is reflected by the targets and certain indicators. For example, targets 7.b and 12.a share the same indicator *“Installed renewable energy-generating capacity in developing countries (in watts per capita)”*. The total number of 12 indicators is repeated through the goals, in 4 cases across three different goals (UNSD, n.d.). More information related to the SDGs,

their targets, but also individual events, publications, actions and the tracking of progress over time can be found at the UN's website related to SDGs (UN, n.d.b).

Further in the thesis the SDGs themselves as well as their targets and indicators are used to connect various general topics to this problematic.

1.3 Online environment, the Internet and Twitter

1.3.1 The Internet and its demography

From its first practical inception in the early 1960s the Internet has gone a long way, from a relatively small-scale network connecting universities and research institutions, it has found itself connecting billions of people who use its merits daily for a wide variety of purposes, from simple things such as news reading and answering emails to complex tasks such as regular content creation in form of videos, blogging platforms, or customized websites for a specific use.

The UN's International Telecommunication Union (ITU), a specialized agency focusing on information and communication technologies (ICTs), has been keeping statistics about the global use of the Internet since the year 2005 (International Telecommunication Union [ITU], n.d.). The rapid spread of the Internet and its services is apparent. While in 2005 16.8% of the people across the globe were using the Internet *regularly* - daily for at least 3 months, in 2010 it was 29.3%, in 2015 41.5%, and finally by 2021 it was 63% of the global population (ITU, 2019; ITU, 2021). The latest ITU report even argues that the global pandemic of COVID-19 has accelerated the spread of Internet usage even further, as for many individuals and companies it had become a necessity during those trying times (ITU, 2021).

However, while the number of 63% might seem impressive, accounting for 4.1 billion people, it is also important to note that the accessibility of the Internet is not uniformly spread throughout the space. Several geographical, social, and economic aspects play an important role when considering such a factor. In the further paragraphs we take a look at some of those factors, which are considered by the ITU document. For the purpose of consistency we use the available data from the year 2020, where 59% of the World's population was using the Internet regularly (ITU, 2021).

Looking into the statistics from a geographical point of view the ITU document considers six macro-regions. We find out that while the numbers for people regularly accessing the Internet are high in Europe (87%), the Commonwealth of Independent States (82%), and The Americas (81%), they get towards the average of 59% in Arab States (66%) and Asia and Pacific (61%), and finally end up way below the average percentage in Africa (33%) (ITU, 2021).

The distribution from the previous paragraph is not random, as we are able to see when we assess each country by its economic strength. The ITU uses three groups - developed, developing and least developed countries. While developed countries have 90% of the population using the Internet regularly, developing countries are at 57%, and the least developed countries at 27%. Yet, while the number for the least developed countries might seem dire it is important to note that while the further spread of the Internet and its technologies within the developed countries has become stagnant, the numbers in developing and least developed countries are still growing rapidly, providing long term potential for especially those populations (ITU, 2021).

We can even judge the situation by the demographics of the Internet users. One such metric, which ties into one of the issues related to SDGs - gender equality, is Internet access considered by gender. Once again, in 2020 59% of the world's population was using the Internet regularly. However, when split by gender, we end up with 62% of males and only 57% of females using the Internet regularly. While overall the gender gap in Internet usage is narrowing over time, especially in the developed countries, it is still present. It is most notable for the least developed countries, with 31% males online and only 19% females online (ITU, 2021).

Similarly, dividing the population into two age groups we can see a clear difference in the overall usage of the Internet. The two groups within the document are the youth and the rest. The youth comprises people aged between 15 and 24 years of age. 71% of people in this age group use the Internet regularly. The other group is people below 15 years of age and over the age of 24. In that age group 57% of the population uses the Internet regularly (ITU, 2021).

Lastly, it is important to distinguish between urban and rural areas. Once again we get two differing numbers. While the global share of Internet users in 2020 for urban areas

was 76%, it was only 39% for the rural areas. Further, while for developed countries the share of urban and rural was 89% and 85% respectively, for developing countries it was 72% and 34%, and for the least developed countries it was 47% and 13% (ITU, 2021).

Based on the previous information we are able to draw certain conclusions about the overall usage of the Internet and its demography for the rest of this work. We can claim that in general the number of people with access to the Internet is rising each year. Further, we are able to determine that people from the urban environment, aged between 15 and 24, and from the regions of Europe, the Americas, or former Soviet republics are the more likely users of the Internet. Further, the economic strength of a region plays a significant role for all the correlated factors, even serving as a magnifier for the metrics which are not as different for the developed countries. Thus, especially for the region of Africa and overall the least developed countries there is a significant gap between the male and female users.

1.3.2 Birth of the social media

Understanding the emergence of social media platforms and individual specifics to each requires a step back in time, to the early days of the *World Wide Web* (WWW), the Internet's most used platform. It was in the early 1990s when the founding stones of the WWW were put together in CERN - Hypertext Markup Language, Hypertext Transfer Protocol, as well as the first browser, first server and first website. This, along with the 1993's release of the WWW technology to the general public has marked the era of the so-called *Web 1.0* (CERN, n.d.). It was an era in which the majority of the online content had been managed by individuals. Era of static websites and simple homepages with predominantly text and inline images, with embedded links leading to other similarly simple websites. The then-users were merely observers, who *consumed* the content available to them, or who exchanged information and files through special online boards (Kaplan & Haenlein, 2010).

The early form of the WWW, the Web 1.0, had not offered many opportunities for user cooperation and communication. However, as the Internet had become more available, the bandwidth and speed had become greater, and the volumes of data transferred through it had increased, a new way of using the WWW emerged. The term *Web 2.0* was first used in 2004 and referred to this new way of utilizing the WWW (Kaplan &

Haenlein, 2010). In other cases, the more descriptive term for such a form of the WWW is used - *participative web* (OECD, 2007). As the second definition hints at, the content was no longer published by individuals on their private homepages, but instead was created in a collaborative fashion, allowing for projects such as Wikipedia to be created. It was not only this re-thinking of the way the Internet is used, but the web browser software had undergone a huge technological improvement through the years, allowing for technologies such as Adobe Flash, or AJAX to be used. The first allowed for animations, audio and video streams to be added to web pages, which allowed for more eye-catching content to be presented to the users. While the other allowed for only some of the website's content to be shown and updated at times, leading to the birth of the dynamic websites which are much more complex and which led to the introduction of things like online advertisements, e-shopping and user-focused content based on their demographics and other influencing factors such as geographical location, or time of day. Thus, the transition to the Web 2.0 allowed for what we now know as social media to emerge (Kaplan & Haenlein, 2010).

The other important component of social media is its reliance on **user-created content** (UCC) (OECD, 2007), sometimes also referred to as **user-generated content** (UGC) (Kaplan & Haenlein, 2010). Both of those terms refer to the same concept. The OECD study defines UGC as: “1) content made publicly available over the Internet, 2) which reflects a certain amount of creative effort, and 3) which is created outside of professional routines and practices” (OECD, 2007, 9). For the purpose of this thesis we note that the first point of the previous definition also includes the social media content which can be considered “semi-public” - located on a publicly available website which requires its users to register and login to view the entirety of content, produce their own content, or use the full functionality of the website. The UGC is the basis for all social media and social networking sites, as it forms the majority of all the content available. The UGC comprises of texts, images, videos, and other formats of the multimedia. It includes original content by users, content linked from secondary sources, but most importantly individual users reacting and interacting with each other – *socializing* in the online environment and forming *networks* of ideas, topics and users.

1.3.3 Twitter, the microblogging social platform

Twitter is one of the world's most popular social networking services. With 436 million monthly users in 2021 Twitter is the 6th most popular social medium focusing on this type of UGC and 15th most popular service focusing on UGC in general (Statista, 2022a). As of 2021 Twitter is most popular in the United States of America with almost 77 million of regular users. Further, the US is followed by Japan with nearly 59 million users, India with 23.6 million users and Brazil with 19 million users (Statista, 2022b).

The core of Twitter's UGC is the *tweets* - short messages posted by its users. Formerly, each tweet was limited to the maximum of 140 characters. It was until November 2017 when they doubled this length to the maximum of 280 characters. However, along with this limitation on text, each tweet may contain addition content – images and photos, animated images, videos, polls and locations from which they were made. Additionally, the text itself may further contain hashtags, user mentions and links to other websites, along with the previews – thumbnails – of those (Twitter, n.d.a).

While user mentions allow you to quickly access the said user's profile by clicking on the name, hashtags are a bit more complex than that. They are not the specific of Twitter, as they are utilized across many social networking websites or websites with UGC on it in general. They are used as key words to mark the overall topic of each tweet. Have you made a tweet about the Kuznets Curve? You could use hashtags like #economy, or #inequality to contextualize your tweet to a broader topic. You can easily navigate through those topics by clicking on them. Then you are able to see the most popular tweets which include that hashtag, as well as follow and find other hashtags. On its home page, Twitter offers a list of trending hashtags – based on their recent popularity, number of reactions and replies. Hashtags are also used to promote certain events, such as #Oscars 2022, or certain social, or political movements, such as #BLM, or #metoo (Twitter, n.d.a). As such, hashtags can topically span from the broad ones, such as #economy, to a very concrete ones, such as #TheSlap referring to an incident at the 2022 Oscar awards, where actor Will Smith slapped his colleague Chris Rock on stage (Grebenyuk, 2022). Overall, hashtags connect various topics which users can access by searching for a hashtag, or clicking on it.

Users on Twitter can interact with each other. If you are interested in the ideas of a certain user, you can *follow* them. Similarly, you can be *followed* by other users. Or you can *block* someone so they cannot see, or interact with your tweets, or someone can block you in this way. Further, user's profile, or particular tweets can be made *private*, so that only the people you follow can see them. When a user logs in they are presented with the home page, where they can see all the recent activity of the users they are following – their original tweets, along with promoted tweets which are a form of advertisement on the website, or some tweets from users you are not following, but who might interested you based on your previous activity (Twitter, n.d.a).

Users may also interact with each other. There is a possibility to send a private message to another user if they have this option enable. However, the main focus on interactions is in the public eye as users can react to other users' tweets.

For this purpose, we refer to the Figure 1, on the next page, which is a screenshot of a tweet made by the WEF. Further content related to a particular tweet is called *thread*, which includes all the tweets which were made as a reply by other users, or the same user. As was mentioned, 280 characters may be in certain cases rather limiting, so users resort to replying to their own original tweet to further explain the details, or provide context for the original tweet. Right below that we can see profile image, full name of the user (World Economic Forum) and the unique user handle (@wef) under which their profile can be found. The small blue icon next to the name marks the user as *verified*, which means that Twitter officials verified that the profile truly belongs to the organization or individual they claim to be. Further is the text of the tweet – including mentions (@realDonaldTrump), hashtags (#wef20), a link (wef.ch/38m8uJp) and an embedded video. Below we can see the information about the tweet – when it was made, from what kind of device or software, and if available (not shown) the location it was made from. Finally, in the lowest part we can see the number of interactions that tweet received – *retweets*, *quote retweets*, and *likes*. The number of *replies* can be seen by scrolling further down – not shown in the picture. There are also buttons for our interaction with the tweet, from the left, reply, retweet and quote retweet, like, and further options. Similarly, above the reply section (not shown) is a text box available for an immediate reply (“Tweet your reply”) (Twitter, n.d.a).

Standardly, Twitter can be accessed from the desktop and mobile devices using various web browsing programs, or a special application. However, Twitter also provides its API for programmers who can use their own programs and tool to access its content (Twitter, n.d.b).



Figure 1: Example of a Tweet (WEF, 2020)

2 Review of available literature

In this chapter we focus on the literature available to the relevant topics. The main sources used for this chapter were the papers, articles and books published and indexed on the on-line academic databases of Google Scholar, Scopus, and the Web of Science. Each topic was thoroughly searched using relevant keywords. The results were sorted by relevancy and processed before being selected. Further, we have assessed the list of citations and references in each publication to identify other potential sources relevant to the topics.

The three overarching topics we focus on in this chapter are: literature focusing on *sustainability and sustainable development* in general, literature focusing on *the SDGs and/or the World Economic Forum*, literature focusing on the *analysis of social media*.

2.1 Sustainability

Sustainability is the topic of our time. It ties closely to the problematic of the environment, economy, and society. It winds itself through many topics of public and academic debates across the globe. We encounter its underlying ideas everyday – recycling, zero-waste, lowering of emissions, reusable accessories, or environment-friendly products.

The importance of the topic of sustainability may be signified by the existence of whole journals dedicated to this problematic. Among the most common is the *Sustainability* (2022), *Sustainability Science* (2022), *Journal of Cleaner Production* (Elsevier, 2022a), and *Journal of Environmental Management* (Elsevier, 2022b).

Further, we were able to identify a vast volume of publications related to this topic. The overall focus of each varies. Several books and articles focus on assessing the whole picture of sustainability, focusing on its roots and outreach across many aspects of our daily lives (Wilkinson et al., 2001; Portney, 2015; Thiele, 2016). These articles often go into the detail about the roles of sustainability within society and how it interacts with the actors from the private sector, public sector and the government. Some attempt to evaluate the current state of knowledge on the topic (Goodland, 1995; Keiner, 2005; Caradonna, 2014; Mensah & Casadevall, 2019), or ponder the exact meaning of the

term and its implication on the public discourse (Ehrenfeld, 2008; Faber et al., 2010; Scoones, 2010).

These and other publications also focus on the concrete subtopics of sustainability. Among those we identify the topics of corporate social responsibility (Elkington, 2006; Rodríguez-Olalla & Avilés-Palacios, 2017; Zenya & Nystad, 2018), sustainability of cities (Marcuse, 1998; Portney, 2015; Cohen, 2017), technological aspects of sustainability (Thiele, 2016), sustainable tourism (Force & Benessaiah, 2018), and stakeholder capitalism (Freeman & Liedtka, 1997; Schwab, 2021).

2.2 Sustainable Development Goals and the World Economic Forum

The problematic of the SDGs is in and of itself embedded into the global politics through the UN. As such, they are often measured and considered on many levels through many institutions. As such we were able to identify two main discourses in the literature. First is the focus on the hopeful message and the opportunity the global projects such as the SDGs (and the preceding MDGs) provide (Sachs, 2012; Ban, 2019). The other is a reserved approach to the overall optimism and a critique of the SDGs, their exact implications and remarks about lacking or insufficient means to measure their progress with accuracy despite the proposed indicators (Hák et al., 2016; Liverman, 2018), or the examination of possible interconnections and contradictions of the SDGs (Le Blanc, 2015; Nilsson et al., 2016). However, the ideas proposed in the previous publications do not doom the SDGs; instead they propose careful and systematic approach with deeper understanding rather than an over-joyed hastiness. Overall, the literature proposes the idea of integration between the individual SDGs – proposing interconnected goals and targets.

The literature focusing on the WEF itself varies. The available documents often include yearly reports by the WEF itself, regarding the activity of their respective platforms. Those reports are better available in its entirety on the WEF's website (WEF, n.d.b). Further, the literature consists of articles considering the results of various reports (Mazanec & Ring, 2011; Waughray, 2011; GG Gap, 2017), or ponder and attempt to describe the nature and potential of the WEF (Graz, 2003; Pigman, 2007).

2.3 Social media and sustainability

When assessing the topic of sustainability within the context of social media and social networks, we have discovered several possible approaches to the problematic. First is the idea of utilizing social networks as a medium to promote companies or individuals through the topics of sustainability (Reilly & Hynan, 2014; Stevens et al., 2016). Further, some focus on the ideas of promoting sustainability itself (Stevens et al., 2016; Hamid et al., 2017), or using the available data for further research (Manetti & Bellucci, 2016; Ilieva & McPhearson, 2018). Finally, some papers see the social media as an overall opportunity for analysis, focusing on topics such as Big Data and sentiment analysis (Kaplan & Haenlein, 2009; Calcagni et al., 2019; Ballestar et al., 2020).

Overall, the literature focuses on the ideas of marketing and promoting of sustainability, as well as the possible opportunities which the social media provide for research and utilization of the broadening options of data processing and analysis.

3 Methodology and goals

3.1 Data formats and technology

While developing custom software to cover all needs for the data processing in this work would be an option, we decided to use some of the already available tools and applications wherever possible. Working with the data requires multiple steps before it can be properly interpreted. Those include - downloading the data in a program-readable format, further reformatting and cleaning up the data, and using the new data as an input for further analysis. Further, statistical and new output data are generated, which are then used for further analysis and to draw conclusions.

For these purposes multiple data formats and programs were used. To be exact, the JSON and CSV formats were used as an intermediary between programs, while the scraping tool Snsrape was used for initial data processing, the spreadsheet program Google Sheets was used for analytical and statistical purposes and the Java programming language was used in between to develop a simple application for more complex calculations and as a mean to remove redundancy from the scraped dataset.

3.1.1 JSON and JSONL

The JSON stands for the *JavaScript Object Notation*, it is a lightweight format based on the notation of the JavaScript programming language used for data interchanging between programs. However, it is independent of any concrete programming language while following the usual conventions of the standard programming languages, such as C, C++, Java, JavaScript, or Python. As such it is an ideal candidate for the data-interchange between various programs, especially those with more complex structure. Further, the aforementioned programming languages often include one or more already developed libraries containing the required functionality for parsing and processing of the JSON files, which simplifies further work.

The premise of JSON is built on two structures: *objects* and *arrays*, and the fact that those structures can be further nested. An object consists of two parts, a pair of *name* and *value*. The name serves as identification for the value, while the value itself may be a text string, number, boolean (true/false value), different object or array, or even null

value. Each object begins with { and ends with }. An array is an ordered collection of values, including all of the aforementioned, which are separated by a comma character. The array itself begins with [and ends with]. Further, the format uses a multitude of special and escape characters to prevent misinterpreting of the data (ex. if we would want to use { as a normal character within a string value) (Introducing JSON, n.d.).

The JSONL stands for the JSON Lines. Where the standard JSON format operates with the whole provided file as one object, the JSONL allows for multiple independent objects which are separated by a line separator character, standardly noted as the ‘\n’ and represented by the ASCII Line Feed character with the value 0x0A. This format allows for multiple objects to be stored within one file, which prevents cluttering of directories by multitude of individual JSON files (Ward, n.d.).

Possible alternatives with similar functionality are the XML, SDL and YAML formats.

The standard file extension for JSON format is .json, and the standard file extension for JSONL format is .jsonl.

3.1.2 CSV

The abbreviation of CSV stands for the *comma-separated values*, which is a blanket term for a text file format. In the CSV format each line of text represents one data record, which is further separated into individual fields, as its title suggests, by commas. It is often used to transfer a multitude of tabular data which all have the same number of fields.

Usually, but not necessarily, the first line of a CSV file includes a header line which includes names of all the fields. Further, depending on the concrete use case of a CSV file, the individual fields are not necessarily separated by a comma, but semicolon, tab, or space may be used. To an extent, any character may be used as a separator depending on the current circumstance. Some more information about the standardization of the CSV format is available in a memo by Shafranovich (2005).

The standard file extension for a CSV file is .csv.

The CSV format was used as an intermediary format multiple times in this work, usually as a mean to copy and paste multi-line and multi-field data to a spreadsheet editor, which includes the functionality to parse imported data by a character. For

example, the information about individual hashtags, mentions, or keywords were extracted along with the number of mentions. Throughout this work, the standard separator used was a space character, with the exception of datasets including multi-word fields where semicolon was used instead. Further, no header lines per se were used. Rather, the first line often included additional info, for example the total number of unique hashtags, mentions, or keywords.

3.1.3 Snsrape

Snsrape is a python-based data scraping application developed by the GitHub user JustAnotherArchivist which is available under the GNU General Public License. It is a multipurpose scraping tool focused on social networking services - SNS, as the first three letters suggest. It allows for multiple data points to be scraped, including information about user's profiles, their posts, or to conduct searches, return results, and explore hashtags. The currently supported social media are the following - Facebook, Instagram, Mastodon, Reddit, Telegram, Twitter, VKontakte, and Weibo.

To run snsrape the Python 3.8 or newer has to be installed on the computer. The application itself has several further dependent libraries to run, but those are installed automatically with snsrape. The application itself is command-line based with no further graphical interface and is run using fixed syntax along with some required and some optional arguments for the program. Among the options, the snsrape allows for the JSONL to be used as a return format.

The generic form of snsrape's command:

```
snsrape [GLOBAL-OPTIONS] SCRAPER-NAME [SCRAPER-OPTIONS] [SCRAPER-ARGUMENTS...]
```

An example of the snsrape usage provided by the GitHub's read-me page:

```
snsrape twitter-user textfiles >twitter-@textfiles
```

The command above returns information about a twitter user operating under the handle *textfiles* and redirects the output to a file with the name *twitter-@textfiles*.

Further information about the usage of the tool can be found on its respective GitHub page, either in the provided read-me document, or within the code itself (JustAnotherArchivist, n.d.).

3.1.4 Java

Java is a high-level object-oriented programming language which thanks to its technology and the way it functions allows for an easy multi-platform use. The only requirement for an application developed in Java to run on a computer, or other devices is to have the *Java Runtime Environment* installed (Oracle, n.d.).

Java was the programming language of choice for the data processing for this thesis because of the author’s long term experience with said language.

The data processing tool itself was developed using the *Eclipse IDE for Java Developers* which is as a software part of the Eclipse Foundation project. The Eclipse IDE provides a useful toolkit for Java developers, along with possibility of integration with online repositories such as GitHub and similar (Eclipse Foundation, n.d.).

In Figure 2, attached below, is a screenshot of the Eclipse IDE – the graphical interface in which the tool used for this thesis was developed (“sample” code can be seen in the middle).

While Java is the programming language of choice in this case, tools for similar purposes may be developed practically in any programming language depending on the concrete purpose of the output data, the concrete format of its input, or simply the author’s familiarity with a certain programming language.

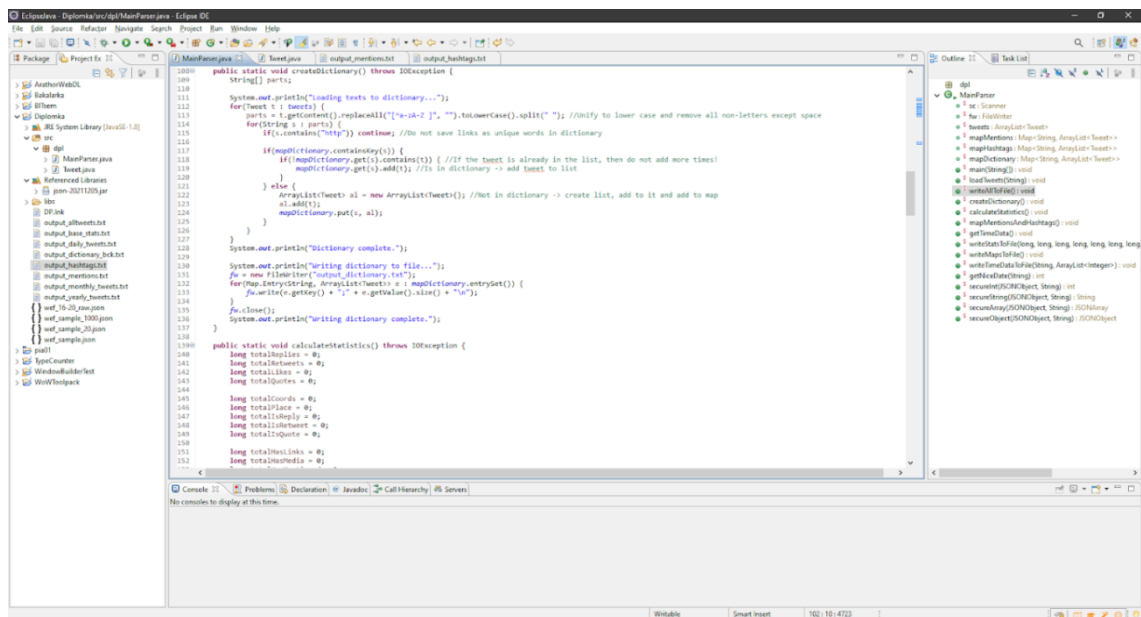


Figure 2: Graphical interface of Eclipse IDE

3.1.5 Google Sheets

Google Sheets is a free spreadsheet program developed by Google. It is a part of their wider online toolkit of Google Docs Editors. It provides online alternatives to the commonly used Windows programs like MS Word and PowerPoint with Google Sheets being the alternative to MS Excel.

Similarly to other spreadsheet programs, Google Sheets include a variety of tools for statistics, analysis and graphical outputs. The program is online and can be accessed through a web browser from anywhere, allowing for the user to edit and work with the same spreadsheet independently of their personal computer. The file is saved to the online cloud platform, accessible through the user's Google Account. Along with the other programs from the Google toolkit it allows for an online cooperation and sharing as well, which is useful for both collaborative work, or to simply getting an (almost) real-time feedback on certain issues, used methodology, or similar (Google, n.d.).

It was chosen as a free and easily accessible alternative to the MS Excel as it provides the similar functionality with the added benefit of an online environment which allows for it to be easily accessed from anywhere and even allows for easy cooperation if need be.

3.1.6 ESRI ArcMap

When working with the data by countries, for some parts of the analysis we used the ESRI ArcGIS desktop software. We used the ArcMap to produce the cartogram presented in this thesis (ESRI, n.d.).

Further, we used the publicly available map data from Natural Earth for the cultural map of the world we used in the ArcMap software (Natural Earth, 2022). The map used in this case was the *Admin 0 – Countries*, to which a few edits was done. Further, the regions were grouped into polygons by the attribute regarding the sovereign dependency. Thus, we ended up with the list of 200 countries and territories which we used further in this thesis.

3.2 Goals

The World Economic Forum is a high impact organization which focuses on the cooperation between private and public sector. They established multiple platforms for the mobilization of the private sector, from individuals to companies, to take action in the contemporary topics of sustainability, technological progress, social equality and others. The main body of their efforts today is based in the online environment through their website and profiles on multiple social media. We believe that analyzing and processing such content may prove insightful from the multitude of possible points of view.

As such, the first goal of this thesis is to describe the presence of the WEF in the online environment with the main focus being the WEF's operation on the social network Twitter. We examine the overall metrics regarding their profile and assess the availability of relevant data.

We then focus on the selected dataset of tweets posted from their profile for the period of 5 years, between January 1, 2016 and December 31, 2020. We analyze the information available in the acquired dataset focusing first on the basic statistics of the WEF's posts – average number of posts, average number of interactions, etc.

As our second goal, we establish three main categories we use for further analysis and processing and which we use to produce relevant outputs based on our data.

First, we establish a connection between the text of a tweet and the possible geographical location it relates to. The selected categorizations were individual countries and some well-established approaches to regionalization of the world. The planned output is the list of countries and the information about posts they were included in. We plan to utilize Twitter's location functionality along with keyword-based searches.

Second, we focus on the understanding of the context of each tweet. For this purpose we utilize the functionality of Twitter – hashtags and user mentions. We plan to list all the hashtags and user mentions by popularity of use. Based on those two metrics we attempt to categorize the tweets into overall topics related to the most popular hashtags and

users. In this case we plan to present the lists of the most popular hashtags and users and connect them to the SDGs problematic.

Third, we focus on the most impactful tweets by the WEF. Those represent the most popular tweets by the number of reactions by other users – like, retweet, reply and quoted retweets. We analyze the text of those tweets to extrapolate the major topics and content related to those. Here we plan to present the important keywords and topics related to each tweet along with the key information. We plan to select the tweets with the highest numbers of relevant reaction type.

4 The World Economic Forum and its online presence

In this chapter we examine the forming ideas behind the WEF along with the current state of the Forum and its presence on the social media with a special emphasis on WEF's Twitter profile.

4.1 History and main ideas

The year 2021 was an important milestone for the World Economic Forum (WEF), since it marked 50 years since its foundation by Klaus Schwab, who is a professor of business and management at the University of Geneva. It was under the name of European Management Forum in January of 1971 when the first annual meeting was held in Davos, Switzerland. The founding idea for the Forum was what Professor Schwab calls "stakeholder capitalism". He proposed the idea that each company should aim to satisfy and serve all of its stakeholders, not just its shareholders. This idea is resonating through all WEF's meetings, projects, platform and partners they cooperate with (World Economic Forum [WEF], n.d.). The relevance of this idea for the WEF can be seen in Prof. Schwab's recent book *Stakeholder capitalism: A global economy that works for progress, people and planet* (Schwab, 2021).

A shareholder of a certain company may be an individual, company, or institution who literally owns a share of the said company via equity stocks and thus has direct financial interest in its profit. On the other hand, a stakeholder is someone who is influenced by a company's success or failure, being affected by the company's presence, or its policies (Schwab, 2021). These ideas were first summarized by Professor Schwab in 1973 in a short document called *The Davos Manifesto* (Schwab, 1973). This document focuses on the main ideas and the code of conduct for managers, highlighting the ideas of serving the society as a whole. In the section marked as B. 4. we may read the following: "The management has to serve society. It must assume the role of a trustee of the material universe for future generations. It has to use the immaterial and material resources at its disposal in an optimal way" (Schwab, 1973). While not worded explicitly, the underlying ideas of social responsibility and sustainability are present in those words.

Professor Schwab further revisited *The Davos Manifesto* in 2020 (Schwab, 2020). In this new iteration he focuses on companies as a whole, not just its managers and he once

again emphasizes the importance of stakeholders and social responsibility. He promotes sustainability, diversity and fairness within the “digital era”. The ideas such as circular and shared economy and innovative technologies are mentioned. He further underlines the non-economic aspects of companies, arguing that their success should not be measured solely by financial returns, but also reflects the environmental and social objectives (Schwab, 2020). These ideas are further mentioned in the contemporary literature on the sustainability of companies and their social responsibility (Rodríguez-Olalla & Avilés-Palacios, 2017; Zenya & Nystad, 2018). Prof. Schwab also mentions the idea of global citizenship - global stakeholders who should strive for cooperation and collaboration for a better world (Schwab, 2020).

Thus, the WEF, through its annual meetings in Davos, more frequent local meetings, or various online activities, seeks to provide a platform for business owners, politicians, academics, or influential individuals to meet with their stakeholders – the public. They aim to be a leading global platform for private and public cooperation (Pigman, 2007).

Since its inception the idea of the WEF was to connect people from all around the globe and to share their ideas internationally. These efforts are seen through over the 50 years of the Forum’s existence. In its first meeting it had invited people from the EU’s European Commission, managers from Europe’s leading companies, as well as the leading academics from the US universities to share their opinions on topics and ideas related to management in companies. Those were the first establishing ideas for the global efforts of the Forum. In the years to follow, the WEF, in cooperation with the UN's Industrial Development Organization (UNIDO), offered a platform to emerging economies of countries such as Bolivia, Nigeria, or Thailand. Soon after, in 1979, the WEF began its long lasting cooperation with the People’s Republic of China. Since then a close and long standing relationship has been formed, which includes the annual events such as “China Business Summit”, or “Annual Meeting of New Champions”, which is colloquially known as the “Summer Davos” (WEF, n.d.e). Over the years the Forum became one of the most influential global institutions with its annual Davos meeting drawing the attention of media and politics alike. High-profile politicians, activists and academics take turns making speeches on its stage. While praised by some and deemed controversial by others, the global role of the WEF in economy, development and globalization is undeniable (Pigman, 2007). Over the years it has

centralized itself in the talks about sustainability, climate action, and environment with a significant online activity to support its idea of private and public cooperation.

4.2 Online presence

The center of the WEF's online presence ties to their official website – www.weforum.org. It offers access to the multitude of content related to its activities and ideas. We are able to explore the summary of their mission, read into promotional materials about the Forum's history and its impact through the years.

Similarly, we are able to access lists of the Forum's partners, platforms and communities along with the other content produced by the WEF – events, reports, videos, podcasts and articles.

The 18 platforms stand at the forefront of their online content – presenting the main topics of interest for the Forum they promote and mobilize its communities into action, to cooperate with organizations and institutions around the globe. Said platforms tie closely to the list of topics featured on the Forum's website. Those topics often coincide with the platforms, focusing on economy, digitalization, climate action, sustainability, social stability, health, AI, and others. Based on those we are able to track and follow related articles, projects and other activities by the Forum. Example of visual representation of the topics can be seen in Figure 3, on the next page. The website's visitors have the option to create an account with which they can log in to the website. That allows them to follow and filter through the various topics and receive content based on those criteria. Further, we are able to follow two links to the more specialized websites – *Strategic Intelligence*, focusing on the exploration and monitoring of the global transformation of economy, industries and issues, and *UpLink*, which is the WEF associated open platform for innovation (WEF, n.d.a).

Finally, the website contains links to their social media accounts, through which the Forum promotes all its other content mentioned above. Depending on the language of the site we can follow the links to their profile on Facebook, Twitter, LinkedIn, Instagram, Flipboard, Tiktok for the English version of content. For the Spanish version we can visit their profile on Facebook, Twitter, LinkedIn, Flipboard, and Instagram. For the Japanese version we can visit their profile on Facebook, LinkedIn and Youtube.

Finally, for the Chinese version of the website we can visit their profiles on Weibo, Wechat, and Ximalaya (WEF, n.d.a). We further focus on the content on the main account of the WEF on Twitter, but we also explore the related profiles which the Forum utilizes for some of their more specific topics and agendas.

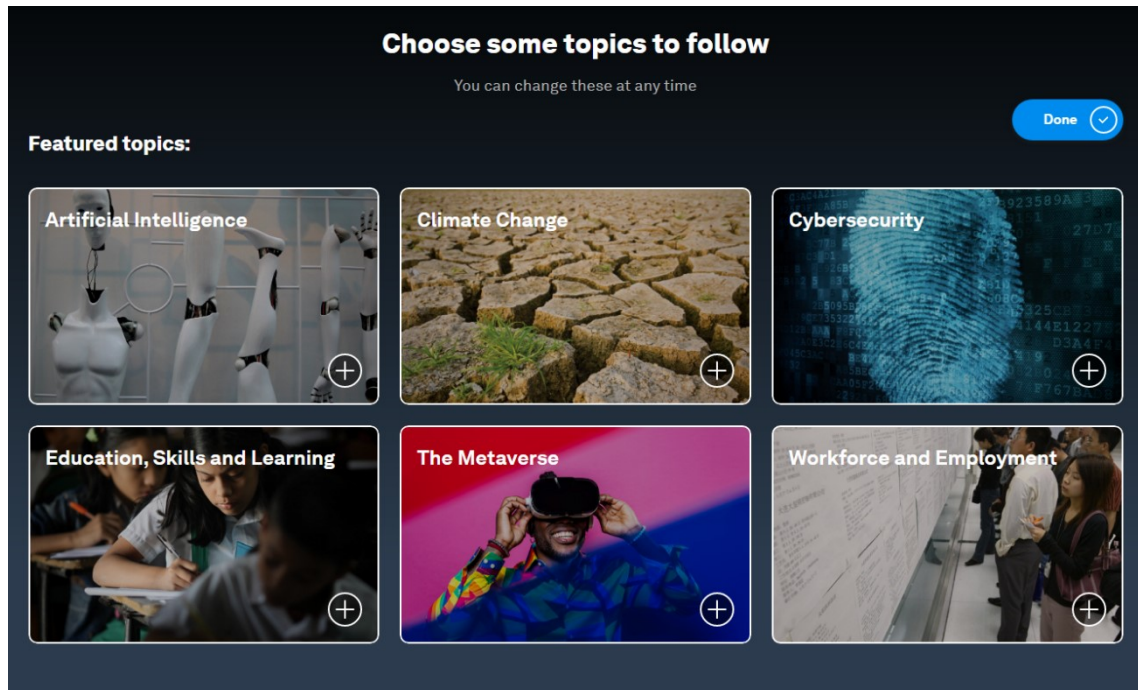


Figure 3: The WEF website, Featured topics (WEF, n.d.a)

4.3 Twitter account

World Economic Forum joined Twitter in April of 2007 under the handle @wef and over the 14 years of its presence on this particular social media platform it amassed over 4.1 million of followers and posted more than 166 thousands tweets. Its profile also contains information about the user's location of Geneva, Switzerland, which is the place where WEF was originally established and is currently a place where the organization's headquarters reside. We also get the link to WEF's official website, www.weforum.org, where the most of its online content is located, including articles, videos, forums and other media. Finally, we can also read the brief description provided about the profile by its user: *"The international organization for public private cooperation. #wef22"* (WEF, n.d.c).

It is also important to note that the language in which WEF tweets is English. While Twitter allows for automated translation of tweets from many languages, there is still

bias towards Anglophonic users. Furthermore, the WEF account operating under the handle of @wef is not the only one managed and related to the WEF itself. From a brief search of Twitter users with the keywords “World Economic Forum” and “WEF” we are able to find at least three more accounts. First one, under the same name World Economic Forum and the handle @Davos is an automated account focusing on retweeting information from the selected group of users. This account has 855.5 thousands of followers (WEF, n.d.d). The other two are WEF Energy with the handle @wefenergy and WEF Cybersecurity with the handle @WEFCybersec, both focusing on topics related to their respective WEF platforms and topics contained in their account’s names. Both of the accounts have less than 5 thousands followers and their activity is much less prolific compared to the main and automated account of the WEF (WEF Energy, n.d.; WEF Cybersecurity, n.d.). However, we further focus solely on the WEF’s main account, as it is the one accessible through their official website, often promoted within their other content, such as articles, and it is the profile with the highest follower count and one promoting primarily their own content and opinions.

While the 4.1 million followers might seem like a large number in Social Blade’s statistics they are ranked as the 1593rd most followed account (Social Blade, n.d.a). Social Blade is a project which focuses on compiling data from multiple online platforms, such as Twitter, YouTube, and Twitch, on multiple data points regarding their users. They keep track of millions of users of said platforms, providing statistics about each of them including rankings, growth, potential revenue, as well as further more detailed statistics (Social Blade, n.d.b). From their statistics on Twitter users we can further examine the top Twitter users to get a better understanding of the popular topics (Social Blade, n.d.c).

Twitter’s most followed account belongs to the 44th President of the United States Barack Obama. He tweets under the handle @BarackObama and has more than 131 million followers. Ranked 2nd to 4th respectively are Justin Bieber (@justinbieber), Katy Perry (@katyperry), and Rihanna (@rihanna). All of whom are musicians and have more than 100 million followers. In 5th place is Cristiano Ronaldo (@cristiano), famous Portugal footballer with more than 98 million followers.

Similarly, 38 of the top 50 accounts on Twitter belong to individuals who consist of celebrities, famous musicians, footballers, actors, or television personalities. Other

notable users from the list might be Elon Musk (@elonmusk), founder of SpaceX and owner and CEO of Tesla, Bill Gates (@BillGates), co-founder and former CEO of Microsoft and philanthropist, and Narendra Modi (@narendramodi), Prime Minister of India.

The 12 of the 50 accounts are profiles related to famous brands and organizations. The most famous among those, in 11th place with 74.5 million followers, is the online video sharing and social media platform YouTube (@YouTube). Others include news channels and newspapers: CNN Breaking News (@cnnbrk), CNN (@cnn), The New York Times (@nytimes), and BBC Breaking News (@bbcbreaking); football clubs, football tournament, and sports channels: Real Madrid CF (@realmadrid), FC Barcelona (@fcbarcelona), UEFA Champions League (@championsleague), SportsCenter (@sportscenter), and ESPN (@espn). Finally, there is Twitter's own official account (@Twitter) and the US's space agency NASA (@nasa).

All of the aforementioned accounts have more than 37 million followers (Social Blade, n.d.c).

5 Data acquisition, description and basic analysis

Since its creation in April 2007 to the November 1, 2021, where the dataset of tweets was downloaded, the main account of the World Economic Forum on Twitter posted more than 167.9 thousands of tweets. Over the selected time period, from 2016 to 2020, the account posted 129.3 thousands of tweets, averaging 70.75 tweets each day. To evaluate and assess the large volume of posts in its entirety manually one by one would be a herculean task and a rather naive approach.

The modern technology comes to a forefront with the utilization of our custom developed program, as well as already established and publicly available tools and software created by others. These include the search functionality of Twitter itself, tools and software which allows for a fast and orderly acquisition, evaluation, transformation and analysis of a large amount of data, or specialized development environments software for creating custom programs. The main focus in this case was the development of a custom light-weight program for data processing using the Java programming language combined with the utilization of a spreadsheet application (ex. MS Excel, Google Sheets) to produce relevant outputs. To acquire the data we used already existing and publicly available tool Snsrape. Similarly, for some of the searches and context analysis we used the search functionality of Twitter. This process is further described in this chapter.

5.1 Data acquisition

Twitter both as a website and as an individual application, and to an extent practically any mainstream web browser, incorporates a functionality which allows its users to search or filter its content, which is a useful initial tool for information gathering. In and of itself Twitter's own *Advanced search* functionality may serve as a great tool for some precise yet narrow queries.

However, to explore the information contained within such a large volume of tweets more deeply or broadly, or to draw any relevant conclusions, it would require many such searches, which could once again return large quantities of tweets and depending on the search even multiple times. Further, this approach would require a certain "support system" which would allow for the results to be easily saved or transcribed for

further research, and analysis. Thus, while the Advanced search functionality was to some degree utilized, some level of automation and computer-processing of data was pivotal to cover the amount of initial input. One such method is called *data scraping* and was used as a primary method in this work, as this approach covers our needs - extracting large amounts of data from a human-readable media for further analysis.

In the following subsections we further describe the functionality and usage of Twitter's Advanced search, as well as the concrete approach to scraping data from Twitter which was used to acquire information within and about tweets of the WEF.

5.1.1 Twitter Advanced search

Both the website and the individual applications for using Twitter on desktop and mobile allows for two "modes" of searching its content. However, to use either of them and to access and browse the majority of Twitter's content it requires you to be registered and logged in onto the website. Then it allows the user to utilize both the "basic" version of search, as well as the Advanced search.

The simple version of search is standardly available within the immediate user interface, either as a box labeled with a "Search" and an icon of a magnifying glass, or by just the icon. From there it allows the user to search the majority of the website's content using words, hashtags, user handles, and the combination of all the former. Further, it allows for the utilization of complex keywords and functions related to Advanced search and it allows for the Advanced search interface to be opened and accessed.

The Advanced search allows for more refined searches to be queried through a user-friendly graphical interface, which is then transformed to a standard search query using keywords and special characters (ex. minus sign to exclude something from a search).

First, in a category called Words, it allows the user to concretize the words of their search. Whether the results should contain tweets with all of the searched words, the exact multi-word phrase, at least one of the searched words, or excluding certain words. As an example, the resulting search query if we were to search for tweets about *dogs and cats*:

- tweet is to contain both words: *dogs cats*,
- tweet is to contain the exact phrase: "*dogs and cats*",
- tweet is to contain at least one: (*dogs OR cats*),

- tweet is to contain only dogs, but no cats: *dogs -cats*.

This functionality can be combined further, ex. containing either the former, or the latter phrase (“*cats and dogs*” OR “*dogs and cats*”). In addition we can include particular hashtags and choose the language in which the tweets should be written.

The second filtering category, labeled Accounts, regards user accounts. The total of three boxes can be utilized. First, the list of accounts from which the tweets were made. Second, the list of accounts to which the tweets were a reply. And third, the list of accounts which were mentioned within the tweets. As an example, the following query (*from:wef*) (*to:wef*) returns all tweets made by the WEF account, which were also replies to the WEF account. This particular search can be used to identify *threads* - multiple follow-up tweets made by the same account as a reply to the original tweet, expanding or explaining its message, usually because of the 280 character constraint.

The third category is simply labeled as Filters. It allows for the user to define whether the results should include tweets which are a reply to other tweets, and if so, if it should include both the original tweets and reply tweets, or just the reply tweets. It also allows for the user to define if the tweets should include a link to another website and if so, whether to show all the tweets, with or without links, or just the ones with links.

The fourth category, labeled Engagement, allows the user to filter the results by a certain minimum amount of interactions with the tweets. They are able to set the required minimal amount of replies, likes, or retweets the tweets should have.

The fifth and the last category is labeled Dates. It allows for the user to define a timeframe for the results, both from and to dates can be used. However, while a minor shortcoming, it does not allow for a more precise time. For example, we cannot refine the search between 12AM to 12PM.

The aforementioned categories can be combined through the Advanced search functionality, or by using the associated keywords within the individual queries. For example, if we wanted to search for the tweets made by the WEF’s account over the selected time period which include the hashtag *sustainability*, have at least 100 likes and does only include original tweets and no replies, the resulting query would be the following: *#sustainability from:wef min_faves:100 until:2020-12-31 since:2016-01-01 -filter:replies*.

Further, the results of a search query can be further filtered by various criteria. Within the user interface the user can choose whether to display tweets from any user, or just the users they follow and whether to display tweets from anywhere, or just locations near them. Similarly, the user can use Search settings to enable or disable sensitive content, which includes various “not safe for work” topics contained in the tweets, such as nudity, and violent and disturbing content, or if the blocked and muted accounts should be excluded from the results. The user can also filter the kind of content which is returned by those searches. There are five categories in total - Top, Latest, People, Photos and Videos.

The Top category refers to the most popular tweets. The general determinant for popularity is the amount of interactions within a certain timeframe. However, the particular functionality of the algorithm is not publicly known and as such is solely in Twitter’s hands. The Latest category includes the most recent tweets, with the most recent shown on top and with the tweets overall being sorted from the most recent to the oldest. The People category lists the individual Twitter users based on the information provided within their profile, or the general affiliation with a certain topic through their own tweets. The final two categories, Photos and Videos, include tweets which contain the respective selected medium - with Videos including videos and Photos including static images, or animated ones, such as GIFs.

Finally, individual search queries can be saved using the Save search button to then be quickly used at a later time.

5.1.2 Data scraping

As was mentioned before, to properly assess the extent of available information within the tweets, some degree of automation and software assistance is required. Given the type of input, graphically engaging content focused on a human user, the technique known as data scraping was utilized to gather, reformat and save the primary input - tweets made by the WEF’s Twitter account from 2016 to 2020.

In general, data scraping is a term used in programming and data sciences which encompasses techniques and programs which focus on reading and processing data which were formerly intended for a human reader. As such those data include a large amount of redundant information, especially in the case of graphic-heavy websites such

as social media. Those are usually written in the markup language HTML with the support of scripting (JavaScript, PHP) or style (CSS) languages, which provide important information about where individual content of the page should be located, what color should be the background, the frames, what should happen when user clicks certain button, text, or other element of the page, and many others. All this information is of no use for the further analysis of the content of each tweet.

In the Figure 4, below, we can see a simplified example of the visual representation of a website compared to its source code, which the web browsing applications (ex. Internet Explorer, MS Edge, Mozilla Firefox) interprets to show the proper visualization. While for the web browsing application the individual HTML tags (ex. <html>, <body>), or the CSS style tags (ex. H1 {color: blue}) are necessary to properly display the page, to the human end user the relevant information is often the content of the website - its title, texts, media and links leading to other pages (connected by arrows). As such, the basic functionality of scraping software can be imagined as a program which goes through the high-redundancy text on the right and which then creates a concise version of the to-human-relevant information on the left. Further example of this process is in the following subchapter.

Thus, the primary purpose of a data scraping software is to extract only the useful and important data from its input, in this case the Twitter website. The output of such software may then vary, from reformatting it into a different human-readable form, or reformatting it into a format which can be easily used by further programs as an input. In this case we had decided to utilize the already existing tool **snsrape**, which is

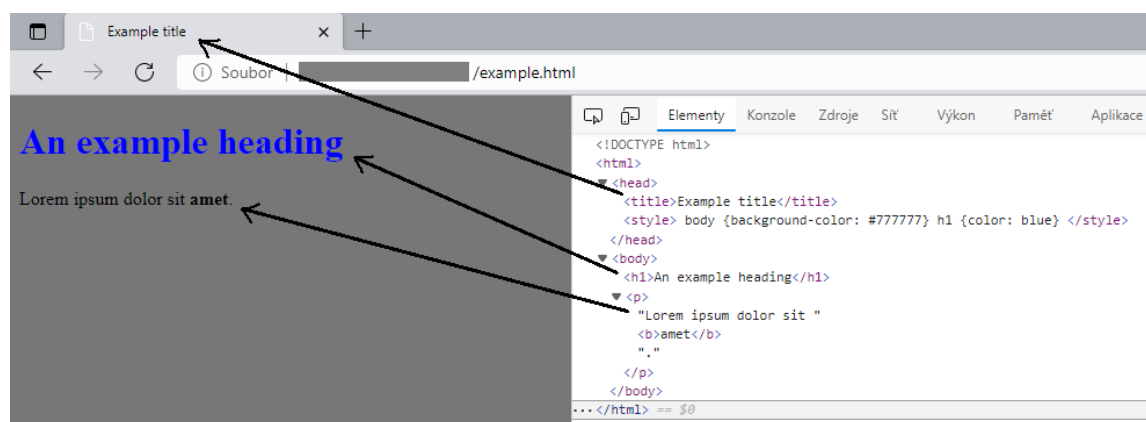


Figure 4 Example of a website's visual representation and its source code

publicly accessible under the GNU General Public License. The data are then used as an input for a program which was custom developed by the authors of this thesis for the purpose of further data gathering and analysis.

The secondary purpose of data scraping, which plays a minor yet still important role for this work, is to create an easily accessible backup for all the tweets over the selected time period. First, while Twitter itself makes certain backups of user data and tweets, those may not be available to all, but to just the user who originally posted said tweets. Further, some of the tweets or the whole profile may be deleted by the user themselves in the future, or it may get blocked, or removed by Twitter for violation of their Terms of Service. Thus, scraping the data and properly archiving them ensures that it will be available in the future for possible future projects and works. Similarly, Twitter is a place of constant interactions among its users. As such the amount of interactions with tweets, followers, blocked users, tweets and many other facets change each second. The data scraping has two impacts on the data: it provides them with certain temporal stability which allows both for drawing relevant conclusions about the dataset, but it may become outdated and inaccurate over time as the online environment changes. However, it also provides an opportunity for further analysis, with data scraped more recently, perhaps even gathered over a different time period and then compared with the original dataset. The other complication with accessing the Twitter data directly is the requirement of almost constant internet connection to access each tweet. Thus, scraping and downloading all the interesting data provides a Twitter-independent and temporally pivoted dataset and backup for this work, further works, or possible cross-validation of results and conclusions.

5.2 Descriptive statistics of WEF Twitter profile

Further we focus on WEF account's prolific activity between the years 2016 and 2020. In the span of 5 years this account posted 129276 tweets averaging 71 posts a day.

Of those tweets only 684 (0.53%) were replies to other tweets, which include replies to WEF's own tweets to create threads related to a certain topic. As an example we can use WEF's tweet regarding Donald Trump's speech at WEF's Annual Meeting 2020 (WEF, 2020). This tweet was followed up by 5 more tweets providing further context to the first one.

Further, the WEF has made a total of 19 (0.01%) quoted retweets and 0 retweets of other users. 46 (0,04%) tweets included spatial information of coordinates and place. However, those had no logical tie to the content of the tweets other than the location from which the tweets were made (24 - Geneve, Switzerland; 18 - South East, England; 3 - Davos, Switzerland; 1 - Paris, France).

Analyzing the content of the tweets revealed further information. The majority of tweets contained both links and some kind of media content, 128129 (99.11%) and 127658 (98,75%) respectively. Of the posts containing media 121052 (94,83%) contained photos, 6067 (4,75%) contained videos and 539 (0,42%) contained gifs (image format used for short animations). Further, 88378 (68,36%) tweets included at least one hashtag and 14432 (11,16%) included at least one mention of another Twitter user.

When analyzing the average amount of interaction with each tweet we come to a number of 70.17 likes, 54.94 retweets, 6.19 quotes, and 3 replies. To put those numbers into perspective, the most popular tweets in each category amounted to 22359 likes, 7554 retweets, 3225 quotes, and 4654 replies.

5.3 Analysis of post frequency

To better understand the behavior and metrics related to the WEF's Twitter profile, it is important to take a look at the posting pattern of their profile. How often they posted a tweet, whether or not that amount changed over time and how did that reflect on other metrics, such as the number of interactions with each tweet, or what might have been the cause for each unusual spike in the number of tweets.

When we break down the number of tweets posted by years for the selected time period, we can come to the following numbers. 35886 (27.76%) tweets posted in 2016, 38976 (30.15%) tweets posted in 2017, 23114 (17.88%) tweets posted in 2018, 15443 (11.95%) tweets posted in 2019, and 15857 (12.26%) tweets posted in 2020. From just these numbers we can see that between the first two years and the last two years the number of tweets dropped to less than a half.

When looking at the data more precisely we come to a number of 70.76 tweets posted each day. This number is, however, far from uniform and we are further able to find local, short term, differences in tweeting patterns, as well as overall global, long term,

differences. As seen in Figure 5, below, showing the number of tweets posted each month, we may recognize two distinct time periods on the long-term scale. Looking into the data more closely, we are able to recognize the 10th April 2018 as the breaking point. Before that date the average number of daily tweets was 102.57 tweets posted each day. After that date the average number of tweets lowered to 44.21 tweets posted each day. The first idea for this change is the change of possible post length, from 140 to 280 characters. However, that change occurred in November 2017 and is thus improbable. We tried to reach out to the WEF via email to receive a clarification for this change. As of April 24, 2022 we did not receive a reply and this question, thus, stays unanswered.

Further, we were able to recognize some of the local extremes in the usual posting pattern of the WEF account as seen in Figure 6, on the next page, showing the number of daily tweets posted as 7 days moving average for better readability. Some of those may be explained by the deeper content analysis of the tweets over those time periods and then tracked back to important meetings or events taking place. As an example, the WEF holds its annual meeting in Davos in middle to late January. Those explain the spikes in the data. However, while in 2016, 2017 and 2018 this pattern included a rise in

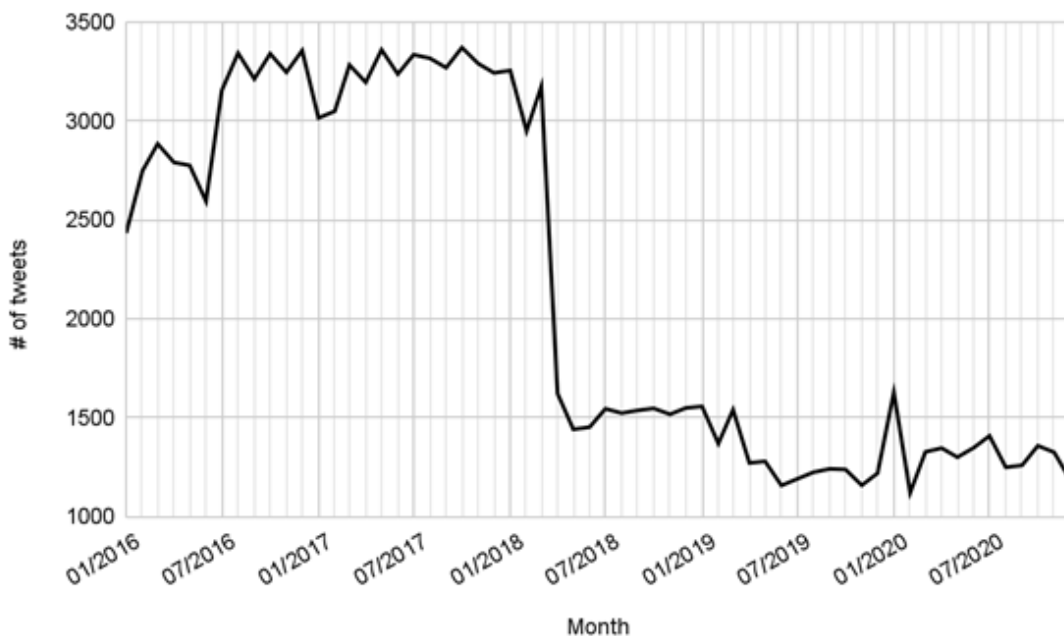


Figure 5: Number of tweets per month

the amount of posts followed by a significant drop, in 2019 we see just a minor change and in 2020 there is only a major upward spike in the number of posts.

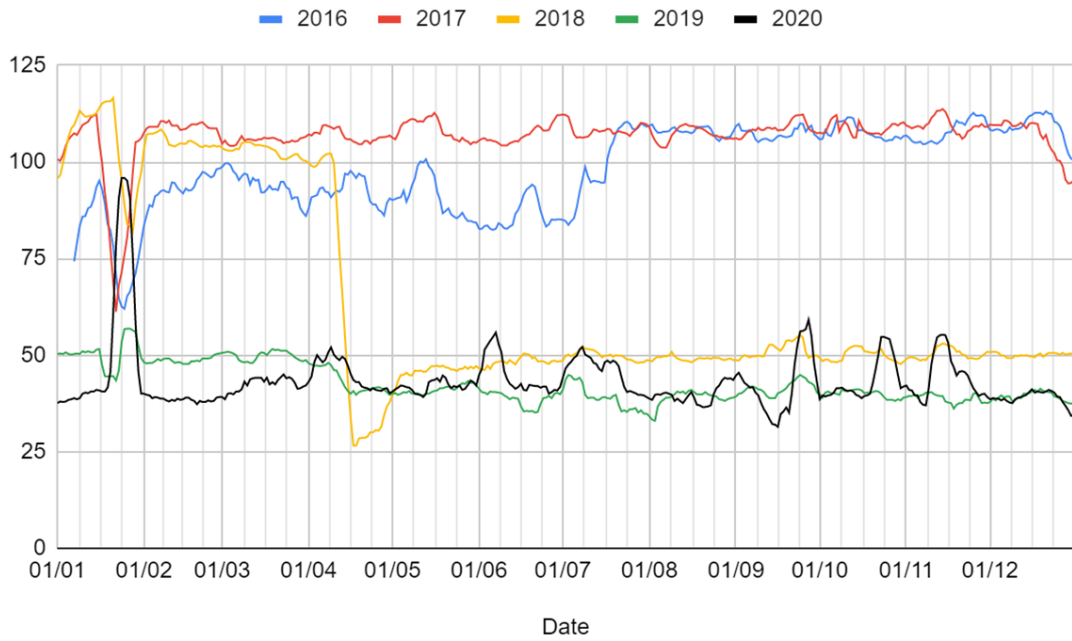


Figure 6: Daily posts by years, 7 days moving average

6 Content analysis and results

In this chapter we focus on the three ways the acquired dataset can be used for further analysis and based on those approaches we present the corresponding results.

The first method we examined more closely is determining the *geographical context* of each tweet. While 46 tweets had information about the place and coordinates from which they were made, those have not proven useful for this kind of analysis. Instead, based on the chosen regionalization a group of keywords was defined for each region. Those were then searched for using the custom software, or the functionality of the Google Sheets to determine the region to which each tweet belongs. We focus on grouping the tweets by individual countries, the World's macroregions - continents, and propose other possible approaches.

The second approach focuses on the overall *topic and context* of tweets based on the inherent functionality of Twitter. That includes the utilization of hashtags as topical labels and the use of user mentions to determine the broader context of individuals and organizations with which the WEF often cooperates. Among those we determine some of the most popular hashtags and users based on the number of tweets they were mentioned in and connect each hashtag to possible relevant SDG topics.

The third approach focuses on the *most popular tweets* among Twitter's user base. For this purpose the four basic types of reactions were considered - like, retweet, reply and retweet with a quote. In each category the top 10 most popular tweets were considered. We further focus on identifying further information - context, topic, related region, included users, individuals or organizations, and related SDGs - and we try to briefly explain the popularity of said tweets.

Finally, while not utilized in this thesis, the three described ways of analyzing the data can be used simultaneously to acquire results with higher levels of detail, or results focused closely on one particular topic - sustainability, European Union, etc.

6.1 Geographical content

In this subchapter we focus on some of the geographical aspects included in the WEF's tweets.

The original text of each tweet was reformatted and cleaned up in a way which did not result in the loss of any information, but which made the searches easier to query. Those adjustments included the removal of non-letter and non-numeric characters and reformatting of the whole text to the lower-case letters only.

Further, a list of broader keywords related to the original keyword was established to acquire much more accurate results about the region's representation. Standardly those included the other part of speech words related to the original term, often including the adjective form related to the original noun. Similarly, the plural forms of the term were considered. As an example, in searches for “Asia” as a region, the words “asia”, “asias”, “asian”, “asians”, and “asiatic” were used. Further, some regions may be, officially or colloquially, known in the English language under multiple names, designations, or abbreviations. For example, searches for “Czechia” as a region included “czechia” and “czech republic” as the base terms and searches for “United Kingdom” included the terms “united kingdom”, “the uk” and “great britain” as those regarding the same region geographically.

While the data themselves allow for a multitude of approaches for determining to which region certain tweets belong to, each of those comes with a certain caveat and requirement for further processing and analysis. We use one such approach, while leaving the others as a possible proposition.

A list of 200 sovereign countries and territories was used as a baseline. The more detailed results are available in the table presented in the Appendix A. Each of those was assigned related keywords and a search was conducted for each of them, providing a number of tweets in which those regions were mentioned. The total of two categories was then examined more closely in the following subchapters - *countries and territories*, and *continents*. Further, we add a third subchapter which discusses other possible approaches to possible classifications.

6.1.1 Countries and territories

In this subchapter we use the total of 200 countries and territories, labeled together as *regions*. The total of 18540 tweets included at least one region mentioned within. That indicates 14.34% of all the tweets. The average number a region was mentioned within a tweet is 92.7 times, with the median number being 6.

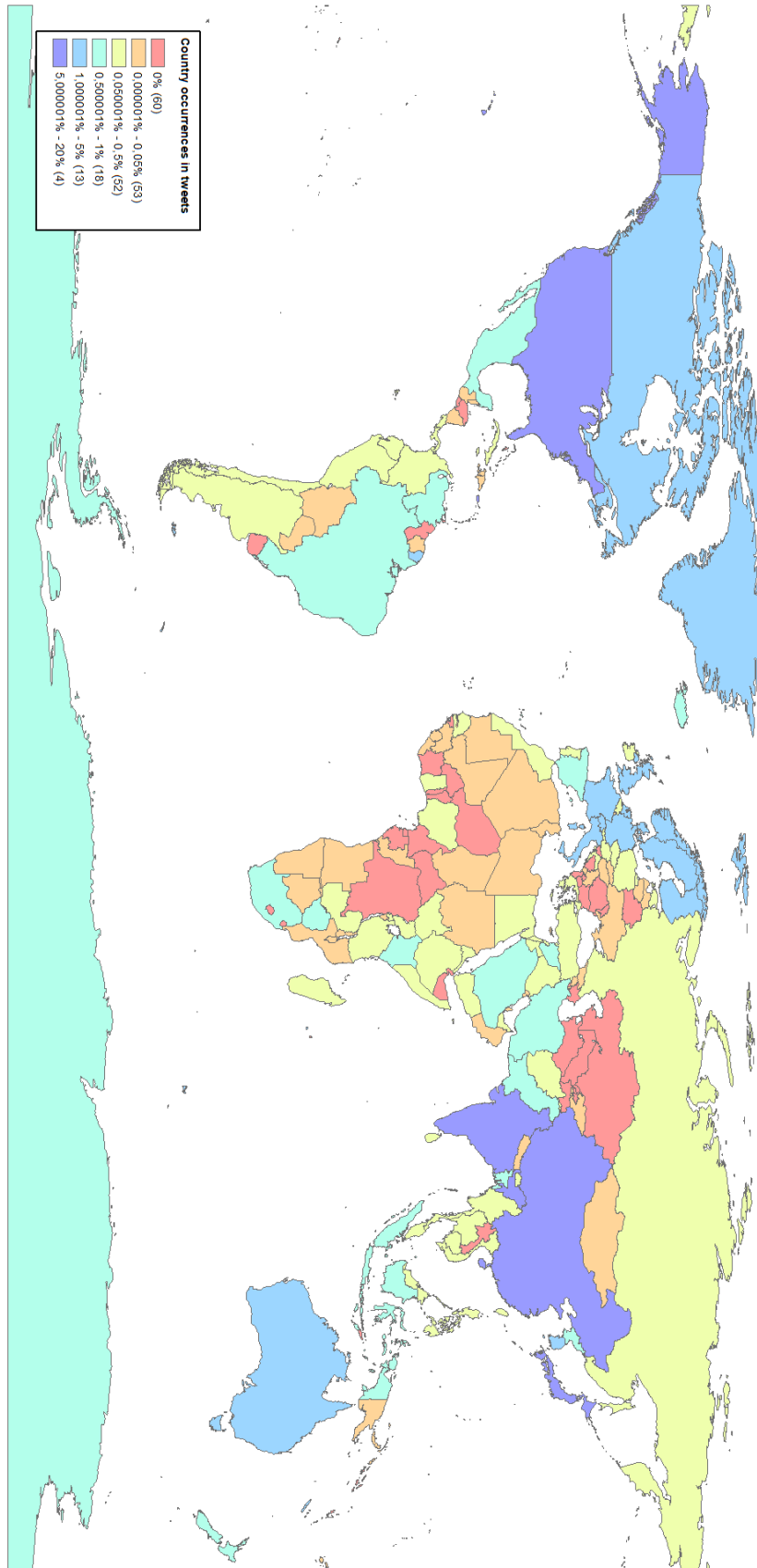


Figure 7: Country occurrences in tweets, % of total (ArcMap)

Those two numbers indicate a high number of low-value occurrences with a low number of extreme high-value occurrences.

Indeed, from the total of 200 regions 140 (70%) of those are mentioned at least once, leaving the 30% countries and territories without a mention. Further, only 85 (42.5%) regions were mentioned at least 10 times, 32 (16%) were mentioned at least 100 times and 4 (2%) were mentioned at least 1000 times. The data are further visualized in the map labeled Figure 7, on the previous page, while the extensive information about each region's occurrence within the data is available in the Appendix A.

For the visual representation a total of 6 groups were used based on the relative number of occurrences. The goal was to use 5 groups where the first group amounts to ~50% of all posts, second along with the first amounts to ~75%, third with the previous amounts to ~90%, fourth with the previous to 99% and fifth with the previous to the 100% of all posts containing a mention. The sixth group contains countries and territories with no mentions.

We were able to identify the top 4 most popular countries and territories, each of which have at least 5% (>927 tweets) of tweets mentioning them and which in total amount to slightly over a half (50.81%) of all the regions mentioned. Those regions were **China** (3430, 18.5%), **India** (2466, 13.3%), **United States of America** (2218, 11.96%), and **Japan** (1307, 7.05%).

In the second group we identified 13 countries and regions, each of which have between 1% (>185.4 tweets) and 5% of tweets mentioning them. Along with the previous group those amount to slightly over three quarters (77.43% of all tweets). Those regions include, among others, and along with the 4 regions from the previous group form the top 10 most popular regions: **Sweden** (564, 3.04%), **United Kingdom** (540, 2.91%), **Germany** (463, 2.5%), **France** (439, 2.37%), **Finland** (417, 2.25%), and **the Netherlands** (351, 1.89%).

Third group includes the total of 18 countries which have between 0.5% (>92.7 tweets) and 1% of tweets mentioning them. This group along with the previous two covers almost 90% (89.86%) of all tweets. Some regions from this group are: **Iceland** (180, 0.97%), **Pakistan** (142, 0.77%), **Saudi Arabia** (130, 0.7%), **Bangladesh** (126, 0.68%), and **Indonesia** (105, 0.57%).

Fourth group includes 53 countries which have between 0.05% (>9.27 tweets) and 0.5% of tweets mentioning them. Along with the previous groups these countries amount to almost 99% (98.81%) of all mentions. Some regions from this group are: **Russia** (85, 0.46%), **Rwanda** (85, 0.46%), **Czechia** (40, 0.21%), **Afghanistan** (25, 0.13%), and **Saint Lucia** (11, 0.06%).

Fifth group includes 52 regions which have at least one mention. Some regions from this group are: **Angola** (8, 0.04%), **Republic of the Congo** (7, 0.04%), **Cyprus** (6, 0.03%), **Chad** (4, 0.02%), and **Kiribati** (1, 0.01%).

The last group consists of 60 regions which were not mentioned within the tweets at all. Some regions from this group are: **Belarus**, **Democratic Republic of the Congo**, **Kazakhstan**, **Uruguay**, and **the Vatican**.

6.1.2 Continents

In this subchapter we use the United Nations geoscheme model which is based on the Standard Country or Area Codes for Statistical Use” (M49) which was introduced by the United Nations Statistics Division (UNSD, 2021). For the purpose of this subchapter only the basic regionalization into the *continental regions* is used. Those include the 5 main continental regions used by the UN to which the independent region of Antarctica is added as the sixth region. The blanket term *continent* is used for all.

Two different approaches for determining the popularity of each region within the tweets were used. First, we use the data for individual countries from the previous subchapter, which we assign accordingly to each respective continent and then draw conclusions from those. Second, as a showcase we use a rather naive approach of using the continent names as keywords. However, such an approach is to be used with caution and serve as a demonstration in this case as the usage of these broader terms as keywords comes with a certain loss of fidelity.

While our main goal in this case is to use the continent names as keywords which relate to a certain place, those are often used in a much broader context. For example, when we were to search for Europe as a keyword, we would use, as previously, related parts of speech forms of the word - “europe”, “european”. However, those would return tweets not strictly related to the certain geographical location, but rather a wide variety

of their usage which often ties with the names organizations, institutions, programmes, projects, and others. Thus, we will end up with results including organizations such as the European Union, which does not necessarily cover the whole region of Europe, or results related to the United States of America, which may not relate to the whole continent of Americas. Thus, further assessment of context would be required in this case to get the accurate representation and in this subchapter serves the mere role of a possible showcase of different approaches to regionalization of tweets according to a certain metric.

The 6 continents which stem from the UN’s M49 classification are the following: *Africa*, *Americas*, *Asia*, *Europe*, and *Oceania* with the addition of the independent region of *Antarctica*. For the first approach, we used the UN’s classification of each individual country or territory from the previous subchapter to assign those to their respective region. For the second approach, we used the continent names as keywords with the addition of the word “australia” for the continent of Oceania. Further, we note that the term used by the WEF to describe that continent is Australia (337 literal results, 516 for individual countries) as the term oceania is not used at all.

In this case we once again operate with the 18540 tweets, as each of the countries or territories belong to a continent, which cover 14.34% of all the tweets. We use the data for individual countries, seen in Table 1, below, and as such the results show the popularity of each continent when considering those - showing the region’s popularity based on the popularity of its countries.

The by far most popular continent is Asia with 9107 uses (49.12%) across 50 countries.

Name	# of uses	% of total	# of uses	# of uses per country
<i>Africa</i>	876	4.72%	(35) 55	15.93
<i>Americas</i>	3243	17.49%	(23) 35	92.66
<i>Antarctica</i>	118	0.64%	(1) 1	118
<i>Asia</i>	9107	49.12%	(40) 50	182.14
<i>Europe</i>	4680	25.25%	(33) 45	104
<i>Oceania</i>	516	2.78%	(8) 14	36.86

Table 1: Continent mentions as sums by countries

The average number of mentions per country is 182.14, where 40 countries (80%) are mentioned at least once. Within the top 5 most popular countries there are 3 countries from Asia - China, India, and Japan. These three alone amass 7203 (38.85%) of all the mentions, while the remaining 1904 mentions (10.27%) are distributed among 37 countries. South Korea is the only other country above the 1%.

In second place is Europe with 4680 uses (25.25%) across 45 countries. Those are mentioned on average 104 times, with 33 countries (73.3%) having at least one mention. While Europe does not have a country among the most popular, with more than 5% of all the mentions, it has 10 countries within the top 18 countries, all of which are mentioned in at least 1% of all tweets. Those countries - from the most popular: Sweden, the United Kingdom, Germany, France, Finland, the Netherlands, Norway, Italy, Denmark, and Switzerland - comprise 3859 (20.81%) of all the mentions, leaving the 821 mentions (4.34%) to the remaining 23 European countries with mentions.

The third more popular continent is the Americas with 3243 tweets (17.49%) across 35 countries. Those are mentioned on average 92.66 times and 23 (65.7%) are mentioned at least once. The by far most popular country from this region are the United States of America with 2218 tweets (11.96%) with 1025 mentions (5.53%) left for the remaining 22 non-zero countries. The only other country from the Americas above 1% of mentions is Canada.

The fourth place is occupied by Africa with 876 mentions (4.72%) across 55 countries. Each country was mentioned on average 15.93 times, 35 (63.6%) of which were mentioned at least once. This makes Africa a continent with the lowest average mentions per country or territory. Some of the countries from this region are: above 0.5% - Zimbabwe (115, 0.62%), and South Africa (99, 0.53%), and above 0.25% - Rwanda (85, 0.46%), Nigeria (56, 0.3%), Ethiopia (50, 0.26%), and Egypt (48, 0.26%).

In fifth place is Oceania with 516 tweets (2.75%) across 14 countries. Each country was mentioned on average 36.86 times and 8 were mentioned at least once. The two most popular countries in this region were Australia itself with 337 (1.82%) mentions and New Zealand with 172 (0.93%) mentions. The other countries were: Papua New Guinea with 2 mentions, and Fiji, Kiribati, Marshall Islands, Nauru, and Palau with 1 mention each.

Concluding our enumeration is the continent of Antarctica with 118 (0.64%) mentions.

Further, we are able to compare this country-based approach with the, albeit naive in this case, approach of using the individual continent names as keywords. We come to the following conclusions, shown in the Table 2, below.

Name	# of uses	% of total	Name	# of uses	% of total
Africa	2410	27.18%	Asia	653	7.36%
Americas	1726	19.47%	Europe	3623	40.86%
Antarctica	118	1.33%	Oceania	337	3.8%

Table 2: Continent mentions by keywords

Overall, these terms were used a total of 8867 times across the 129276 tweets, which amounts to 6.86% of all tweets. The most popular term was Europe with 3623 uses (40.86%), followed by Africa with 2410 uses (27.18%), and Americas with 1726 uses (19.47%). On the lower end of this categorization we have Asia with 653 uses (7.36%), Oceania with 337 uses (3.8%) and Antarctica with 118 uses (1.33%).

From this simple comparison we are able to make a few observations. Africa is the only continent where its individual keyword was mentioned more times than the sum of its countries. Then, on the contrary, while Asia was the most popular of all the continents (49.12%) by individual countries, in this case it is in the 4th place (7.36%). However, it is once again important to note that to assess these data more accurately it would require further data processing to determine the real context of the keywords.

Finally, approaching the popularity of each region by the number of tweets as a metric is but one possible approach. Tweets identified in this subchapter may serve as a future input for further analysis as is shown in the following subchapters 8.3 and 8.4, where these techniques and approaches may be combined to identify organizations, institutions, and individuals often related to certain regions or countries, to determine popular topics of tweets based on hashtags and overall keywords, or to identify the tweets from certain regions with most reactions.

6.1.3 Other possible classifications

Given the availability of the data by countries and territories, we may use those further in a broader context, similar to the approach in the previous subchapter. The country-

based information may then be used further by other classifications, or by using the United Nations M49 geocheme model with more detailed regions - subregional, or sub-subregional classification.

As an example, we use the *sub-subregional* classification which comprises 22 regions in total. The geographical representation of each region can be seen in the map shown in the Figure 8, below.



Figure 8: M49 UNSD Sub-subregions (Wikipedia, n.d.)

Similar to using the continental regions we merge together data from individual countries and regions, based on the sub-subregion they belong to obtain the results. The outcome for each region is presented in the Table 3, on the next page.

To further fill in this distribution, we may once again use the names of each respective region as a keyword. However, for the total of 22 regions only 10 of those were mentioned as keywords in 188 tweets (0.15%).

Those regions are Western Europe (44, 23.4%), South-eastern Asia (43, 22.87%), Eastern Europe (27, 14.36%), South America (27, 14.36%), Caribbean (24, 12.77%), Southern Africa (12, 6.38%), Australia and New Zealand (6, 3.19%), Central Asia (2, 1.06%), Western Asia (2, 1.06%), and Eastern Africa (1, 0.53%).

Geoscheme region	Mentions	% of total	Geoscheme region	Mentions	% of total
<i>Australia & New Zealand</i>	509	2,76%	<i>Northern America</i>	2502	13,58%
<i>Caribbean</i>	71	0,39%	<i>Northern Europe</i>	2406	13,06%
<i>Central America</i>	205	1,11%	<i>Polynesia</i>	0	0,00%
<i>Central Asia</i>	8	0,04%	<i>South America</i>	465	2,52%
<i>Eastern Africa</i>	538	2,92%	<i>Southeastern Asia</i>	471	2,56%
<i>Eastern Asia</i>	5203	28,24%	<i>Southern Africa</i>	107	0,58%
<i>Eastern Europe</i>	170	0,92%	<i>Southern Asia</i>	2897	15,73%
<i>Melanesia</i>	3	0,02%	<i>Southern Europe</i>	580	3,15%
<i>Micronesia</i>	4	0,02%	<i>Western Africa</i>	125	0,68%
<i>Middle Africa</i>	19	0,10%	<i>Western Asia</i>	528	2,87%
<i>Northern Africa</i>	87	0,47%	<i>Western Europe</i>	1524	8,27%

Table 3: Sub-subregion mentions by countries

The proposed idea of using the countries as a baseline for the measuring of popularity of regions can be extrapolated to other metrics and categorizations beyond the one above. It is, however, advised to use those along with the group of keywords suitably selected for the chosen categorization.

Some examples of the possible approaches may be the utilization of OECD regionalization based on the country's GDP, focus on important global, or regional organizations (NATO, African Union, G7, ASEAN), or commonly used academic, or business designations for certain regions (EMEA - Europe, the Middle East and Africa, MENA - Middle East and North Africa), and others.

Finally, dividing the WEF's into geographical regions of choice may serve as a first step for a deeper analysis and research, which may further focus on other metrics and information which are discussed in the following subchapters. Those are the important actors for those regions - individuals, organizations, businesses and others, or important topics discussed - based on the analysis of hashtags, or other keywords. Similarly, based on the number of reactions, important and influential tweets for each region can be

identified to better understand users' sentiments about concrete problematic from topical or geographical point of view.

6.2 Topical and context approach

In this subchapter we focus on the two inherent ways a user can indicate a topic of a post, or contextualize a post on Twitter. The first way to do so is through the use of *hashtags*. The other way is the utilization of *user mentions* within tweets. Both of those can be used as a reliable metric for tracking the topic and context of individual tweets.

In the following subchapters we focus on each of these approaches individually in the broad context of the whole dataset.

6.2.1 Popular hashtags

From the number of 129276 tweets in total the 88378 (68.36%) of them included at least one hashtag. Overall, 2002 different hashtags were used a total of 126078 times. The average number a particular hashtag was used is 62.98.

For the further analysis the total number of 27 most popular hashtags were chosen, shown in the Table 4, on the next page, where the boundary was at least a thousand uses of that particular hashtag. The 27 most popular hashtags amount to the total of 69113 uses, which contains 54.8% of all hashtags. The most popular hashtag, *leadership*, was used 8585 times, while the least popular hashtag within the group, *edchat*, was used 1026 times. Further information about popular hashtags can be seen in the table below.

Further, we are able to identify important topics, events, organizations or institutions to which the hashtags refer and connect them to a broader and similar topic, to individual SDGs, or to a particular geographical location.

First, among the 27 most popular hashtags we are able to recognize 5 hashtags which relate to concrete geographical regions. Two of those regard whole continents - Europe and Africa. The other three regard countries, which are at the same time the most populous countries in the World - China, United States of America and India.

Rank	Name	Count	R.	Name	Count	R.	Name	Count
1	<i>leadership</i>	8585	10	<i>gender</i>	2430	19	<i>climate</i>	1291
2	<i>technology</i>	6233	11	<i>cities</i>	1979	20	<i>bestof</i>	1238
3	<i>environment</i>	5952	12	<i>energy</i>	1827	21	<i>africa</i>	1232
4	<i>health</i>	5404	13	<i>ai</i>	1824	22	<i>ageing</i>	1110
5	<i>economics</i>	5180	14	<i>china</i>	1622	23	<i>innovation</i>	1093
6	<i>education</i>	3612	15	<i>society</i>	1521	24	<i>wef17</i>	1088
7	<i>covid19</i>	3166	16	<i>europe</i>	1443	25	<i>space</i>	1031
8	<i>work</i>	2823	17	<i>us</i>	1430	26	<i>india</i>	1028
9	<i>coronavirus</i>	2617	18	<i>climatechange</i>	1328	27	<i>edchat</i>	1026

Table 4: A list of the most popular hashtags

Further, 3 hashtags include events, or broad thematic topics. The hashtag *wef17* is used by WEF to label tweets which are connected to their annual Davos meeting, in this case in 2017. While this one is the most popular among others (1088 uses), similar hashtags were used to mark tweets for Davos 2015 (2), 2016 (3), 2018 (856), 2019 (884), 2020 (827) and the upcoming Davos 2021 (108). The other two frequently used hashtags were *bestof* and *edchat*. While the first one focuses on lists denoting the best in certain categories, among many: health, pollution, education, corruption, safety, economy, society, and environment. The latter focuses on education and learning with brief informative diagrams, or videos on a wide variety of topics. Similar to the *wef17* hashtag the *bestof* has more iterations - *bestof2017* (384), focusing on year 2017, *bestofdavos* (70), focusing on the annual Davos meetings, and *bestofamnc* (29), focusing on the Annual Meeting of the New Champions which is an event held each year by WEF's in China. The *edchat* hashtag can then be connected to a broader topic of education and the SDG 4: Quality education.

In total 2 hashtags focuses on the ongoing global pandemic of COVID-19 coronavirus disease. First of those is the *covid19* and the other the more general term of *coronavirus*. Both of those can be connected to a broader topic of global health, as well as the SDG 3: Good health and well-being.

We are then left with 17 hashtags which focus on broader, often globally discussed topics. Many of those can be further connected to individual topics of sustainability, or concrete SDGs. We are even able to find the three pillars of sustainable development among the popular hashtags - environmental (environment), economic (economics), and social (society). Similarly, we are able to find topics which focus on the current technological progress predominantly within the field of automation, artificial intelligence and the so called Internet of Things. The often used term for contemporary progress is the 4th Industrial Revolution, which in and of itself is among the popular hashtags on the 28th position as *4ir* with 957 tweets mentioning it. The other related hashtags from the top 27 are - leadership, technology, economics, cities, ai, innovation and to some degree ageing and space.

The further connections between the most used hashtags and the individual SDGs can be seen in the Table 5, on the next page. However, it is important to note that in this case only the dominant subtopics, based on keywords, were considered before the SDG and the hashtag were connected. As such, in future works the popular hashtags could be broken down further into sub-groups to find more precise connections and categories. Similarly, the last SDG, SDG 17: Partnership for the goals, was omitted as it had permeated through all of the topics, with the WEF operating on an international level, in and of itself focusing on global cooperation and partnership. Thus many tweets focus on the international, interregional, or global aspect of a certain issue.

6.2.2 Popular user mentions

From the number of 129276 tweets in total the 14432 (11.16%) of them included at least one user mention. In total, 2727 different users were mentioned within the WEF's tweets amounting to 20637 total mentions with the average amount of mentions per user being 7.57.

For further analysis the top 20 mentioned users were chosen, shown in the Table 6, on the page 67. Those amount to the total of 4342 mentions, which is 21% of all the mentions. The most popular user mentioned is *BillGates* with 582 mentions, while the 20th most popular user is *schwabfound* with 113 mentions. Further information about the most mentioned users can be found in the table below.

SDG	Related hashtags
1 No poverty	(2) economics, work
2 Zero hunger	(0) -
3 Good health and well-being	(3) health, covid19, coronavirus
4 Quality education	(3) education, society, edchat
5 Gender equality	(4) education, work, gender, society
6 Clean water and sanitation	(2) health, cities
7 Affordable and clean energy	(3) cities, energy, innovation
8 Decent work and economic growth	(6) leadership, economics, work, cities, society, space
9 Industry, innovation and infrastructure	(7) leadership, technology, economics, cities, ai, ageing, innovation, space
10 Reduced inequalities	(4) technology, economics, society, ageing
11 Sustainable cities and communities	(4) cities, climatechange, climate, innovation
12 Responsible consumption and production	(7) leadership, economics, cities, energy, ai, climatechange, climate
13 Climate action	(4) environment, cities, climatechange, climate
14 Life below water	(3) environment, climatechange, climate
15 Life on land	(3) environment, climatechange, climate
16 Peace, justice and strong inst.	(1) society
17 Partnership for the goals	- not rated / all -

Table 5: SDGs and related hashtags

Rank	Handle	Twitter Name	Mentions
1	<i>BillGates</i>	Bill Gates	582
2	<i>WEFBookClub</i>	WEFBookClub	555
3	<i>IMFNews</i>	IMF	271
4	<i>StatistaCharts</i>	Statista	270
5	<i>NASA</i>	NASA	255
6	<i>WHO</i>	World Health Organization (WHO)	234
7	<i>YGLvoices</i>	Young Global Leaders	205
8	<i>Google</i>	Google	196
9	<i>amonck</i>	αδριαη	193
10	<i>zahidi</i>	Saadia Zahidi	184
11	<i>UN</i>	United Nations	179
12	<i>WWF</i>	WWF	159
13	<i>wef</i>	World Economic Forum	159
14	<i>Lagarde</i>	Christie Lagarde	159
15	<i>IBGC_Fletcher</i>	The Institute for Business in the Global Context	136
16	<i>JosephEStiglitz</i>	Joseph E. Stiglitz	129
17	<i>WEFUpLink</i>	UpLink	125
18	<i>borgebrende</i>	Børge Brende	121
19	<i>RobinPomeroyEd</i>	Robin Pomeroy	117
20	<i>schwabfound</i>	Schwab Foundation for Social Entrepreneurship	113

Table 6: A list of the most popular mentions

Analyzing the content and context of the selected group of the most popular mentioned users, we are able to identify 3 distinct types of users. First, users who are aligned with or who work for the WEF with three subgroups: *the WEF*, *organizations and projects related to the WEF*, and *individuals related to the WEF*. Second, other users who either cooperate with the WEF or who are important in their respective field with two subgroups: *organizations and projects*, and *individuals*. Third are users related to companies whose main focus is *data processing, statistics and data visualization*.

Within the first group we have assigned a total of 9 users, 1 in the first subgroup, 4 in the second subgroup, and 4 in the third subgroup.

The *World Economic Forum* within the first subgroup, which contains mainly informative tweets with self-mentions, or which are part of a thread.

Further, within the second subgroup there are 4 users. The *WEF Book Club* which focuses on informative and educational tweets including book promotion, based on the current topic of interest (peace, climate action, economic growth, etc.), and on various informative podcasts and similar online content. Then it is the WEF's *Young Global Leaders* initiative which focuses on young talent from all faucets of skills, from art to management, promoting inclusivity and global cooperation for success. The *WEF UpLink* is an innovation platform which focuses on rethinking the way we approach both our work and free time, with main focus on clean energy, climate change, sustainability and responsible consumption and production including recycling and circular economy. Lastly, the *Schwab Foundation for Social Entrepreneurship* is a sister organization of the WEF which focuses mainly on topics related to social issues and social innovation, for example affordable and accessible healthcare, well-being, reduction of poverty and hunger, and reduction of inequalities.

Within the third subgroup there are 4 users. Those include *Børge Brende*, current President of the WEF, former foreign-trade and industry and environment minister of Norway and Managing Director of the Red Cross. *Saadia Zahidi* and *Adrian Monck*, who both work as Managing Directors at the WEF, creating content related to various topics, with *Zahidi's* focus being gender equality, inclusion, and workplace environment, and *Monck's* focus being globalization, technology and the 4th Industrial Revolution and AI. Finally, *Robin Pomeroy* who works as a Podcast Editor for the WEF and who hosts a podcast *Radio Davos*, which can be further tracked under the hashtag *#RadioDavos*.

Within the second group we have assigned 9 users, 6 in the first subgroup, and 3 in the second subgroup.

Among the six from the first group are 4 internationally operating organizations, *International Monetary Fund (IMF)*, *World Health Organization (WHO)*, *United Nations (UN)*, and *World Wide Fund for Nature (WWF, World Wildlife Fund)*, 1

transnational company, *Google*, and 1 independent agency of the US government, *National Aeronautics and Space Administration* (NASA). These are often connected to the WEF through one of their meetings, especially the annual Davos meeting, as members of the aforementioned organizations and companies are often invited as speakers, or panelists for such occasions. Similarly, they collaborate with the WEF on certain projects, or their individual projects are promoted as they often overlap topically. The common topics discussed and mentioned are technology and innovation, and the call for global cooperation and for strong institutions. Further, IMF focuses mainly on economic growth, sustainability and Covid-19 pandemic, NASA focuses on women empowerment, equality and climate change, WHO focuses on health, Covid-19 pandemic, affordable vaccines, and mental health especially at work, Google focuses on technology, innovation, computers and AI, and is often connected to its CEO Sundar Pichai (@SundarPichai, 31 mentions), UN focuses on global cooperation, strong institutions, equality, and climate change and sustainability, and finally WWF focuses on sustainable ecosystems, climate change, wildlife protection and sustainable cities and communities.

In the second subgroup there are 3 users. *Bill Gates*, former CEO of Microsoft, entrepreneur and billionaire who now focuses on foundation and nonprofit work propagating issues related to, but not only, SDGs through *Bill & Melinda Gates Foundation* (@gatesfoundation, 91 mentions). The most common topics are technology and innovation, climate change, zero hunger and zero poverty, and reduction of inequalities. Then it is *Christine Lagarde*, President of the *European Central Bank* (@ecb, 7 mentions) with the dominant topics being women empowerment, climate action, economy and economic growth, and decent work and workspace. Lastly, *Joseph E. Stiglitz*, American economist, professor and holds Nobel Memorial Prize for Economic Sciences, who is mainly mentioned in connection with overall economic-related topics, for example globalization and anti-globalization, capitalism and markets, and GDP as a metric.

Finally, in the third group we have assigned 2 users. *Statista* is a company which focuses on providing statistical data from various industries to its users, focusing on visualization of the output through graphs and infographics and which serves as an often mentioned source for data within some images attached to the WEF's tweets. As such,

the company's Twitter account is often mentioned within those tweets. The other being *Institute for Business in the Global Context*, aligned with *The Fletcher School* (@FletcherSchool, 0 mentions) and *Tufts University* (@TuftsUniversity, 0 mentions), who similar to the company Statista provides publicly available statistical data and infographics which the WEF then uses in their own tweets.

6.2.3 Other interesting hashtags or mentions

In this brief subchapter we enlist some of the hashtags and mentions which were not among the most popular, but still relate to an interesting topic and amassed enough uses not to be omitted.

Some of the other hashtags include, those related to - *technology*: 4ir, automation, science, neuroscience, internet, brain, tech, and robots; *economy*: employment, business, trade, economy, jobsreset, globalization, productivity, finance; *society*: inequality, mentalhealth, migration, transport, happiness, gendergap, socialmedia, psychology, poverty, and demographics; and *environment*: sustainability, nature, food, agriculture, oceans, tourism, water. All of the aforementioned were included in at least 250 individual tweets.

Further, two hashtags relate to a concrete country: japan (481 uses) and brexit (445 uses) which relates to the UK.

Finally, hashtags explicitly mentioning the SDGs were identified: sdgs (225 uses), sdg (7), sdg4 (5), and sdg14 (2).

Similarly with the further mentions. We can identify organizations, companies and individuals aligned, or working at the WEF: *Apolitical* (@apoliticalco) *Strategic Intelligence (WEF)* (@WEF_Intel), *Avery Blank* (@BlankAvery), *Ida Auken* (@IdaAuken), and *Friends of Ocean Action* (@FriendsofOcean). Other organizations, companies and individuals: *OECD* (@OECD), *UNICEF* (@UNICEF), *Harvard University* (@Harvard), *LinkedIn* (@LinkedIn), *Oxfam International* (@Oxfam), *PwC* (@PwC), *Winnie Byanyima* (@Winnie_Byanyima), and *Adam Grant* (@AdamMGrant). And users which are mentioned as a platform through which the WEF communicate some of their content, or which data they use: *Spotify* (@Spotify), and *Apple Podcasts* (@Apple Podcasts).

6.3 Most popular tweets

In this part we focus on the top 10 most popular tweets by all four direct metrics of interaction - like, retweet, reply and quote retweet. For each tweet we focus on identifying its main topic, involved actors and users, connection to concrete SDGs, and potential geographical pivot point. Certain overlaps in the most popular tweets are expected. While each category is treated the same way, the overlapping tweets are mentioned by each respective subchapter.

6.3.1 Top 10 liked tweets

Before we delve further into the tweets themselves, we note that over the selected time period the average amount of likes WEF's tweet had received was 70. Further, the highest number of likes is 22359, 10th highest is 7229 with 14794 average likes in the top 10. The tweets can be seen in Table 7, below.

Keywords	Date	Media	SDGs	Likes
Donald Trump, speech, Davos 2020, economy, USA	21/01/2020	Link, Video	8, 9, 17	22359
GDP, economy report, India	22/02/2020	Link, Image	8	17184
BTS, South Korea, music, globalization	18/12/2018	Link, Image	9, 17	14945
Greta Thunberg, speech, activism, climate change, Davos 2020	21/01/2020	Link, Video	12, 13	11759
BTS, South Korea, music, globalization	01/03/2019	Link, Image	9, 17	11383
Technology, innovation, reforestation, 1t.org	20/11/2020	Link, Video	9, 13, 15	10191
Imran Khan, speech, Davos 2020, justice, Pakistan	22/01/2020	Link, Video	16, 17	8380
Imran Khan, climate action, Pakistan	25/11/2020	Link, Image	13	8122
BTS, South Korea, music, globalization	21/12/2018	Link, Image	9, 17	7919
Top 10, tech industry, USA, Canada, Australia	28/11/2017	Link, Image	8, 9	7229

Table 7: Top 10 liked tweets

In this category, we may find various categories and topics represented in the top 10 tweets. It is important to note that the three tweets (3, 5, 9) which mention the South Korean boy band BTS focus on the same topic. Moreover, two of those tweets are identical and the third, and the most popular, differs just by a missing mention of BTS' official account and exclusion of the #BTSARMY hashtag. We also see a reforestation project led by the WEF 1t.org (6, link: www.1t.org) which focuses on the mobilization of the actors from the private sector to plant over 1 trillion trees by 2030, stemming their goal from the UN's Decade on Ecosystem Restoration.

Three of those (1, 4, 7) are speeches or messages from the WEF's annual Davos meeting from 2020. Three include politicians, Donald Trump (1), at the time US President, and twice Imran Khan (7, 8), the current PM of Pakistan. One includes climate activist Greta Thunberg (4). Three mention South Korean boy band BTS (3, 5, 9). One focuses on educating and informing about reforestation efforts (6). Two focus on listing progress, or best-at-their field - one informing about India's economy (2) and one listing popular tech companies (10).

We are able to identify topics mentioned in the tweets, and discussed in the media attached to those tweets. Further, we are able to link those tweets to particular SDGs, in the table above. The topics which are often mentioned are the following: economic growth (1, 2), culture (3, 5, 9), just and peaceful nations (7), climate change and climate action (4, 6, 8), technology and innovation (6, 10), and global cooperation and globalization (1, 3, 4, 5, 6, 8, 9, 10).

Further, although some of the topics have global outreach, we are able to find concrete geographical locations, in this instance the countries, from which certain topics or ideas stem. Three tweets are connected to South Korea and its musical culture (3, 5, 9). Three are connected to Switzerland's Davos and WEF's 2020 conference (1, 4, 7). Two tweets relate to the USA (1, 10), and Pakistan (7, 8). Finally, India (2), Sweden (4), Canada (10) and Australia (10) are mentioned or can be related to one of the tweets.

6.3.2 Top 10 retweeted tweets

The average number of retweets per tweet for the WEF's account was 55 retweets. Their most retweeted post achieved 7554 retweets, the 10th most retweeted post achieved 2934 retweets. The average amount of retweets for the top 10 is 4488 retweets.

It is important to note that there is a certain overlap between categories. In this instance the total of 6 tweets (1, 2, 3, 5, 6, 9) is also contained in the top liked subchapter. The tweets can be seen in Table 8, below.

Keywords	Date	Media	SDGs	Retweets
GDP, economy report, India	22/02/2020	Link, Image	8	7554
BTS, South Korea, music, globalization	18/12/2018	Link, Image	8, 17	6087
Donald Trump, speech, Davos 2020, economy, USA	21/01/2020	Link, Video	8, 9, 17	5183
Climate action, reforestation, Pakistan	02/07/2018	Link, Video	11, 13, 15	4844
BTS, South Korea, music, globalization	01/03/2019	Link, Image	8, 17	4549
Greta Thunberg, speech, activism, climate change, Davos 2020	21/01/2020	Link, Video	12, 13	4295
Plastic, recycling, circular economy, road infrastructure, India	12/09/2017	Link, Video	9, 12	3451
Joko Widodo, Indonesia, cooperation, economic growth, peace	12/09/2018	Link, Video	8, 13, 16, 17	2996
Imran Khan, speech, Davos 2020, justice, Pakistan	22/01/2020	Link, Video	16, 17	2986
Stephen Hawking, the UK, cooperation, technology, innovation	14/03/2018	Link, Video	8, 9, 10, 16, 17	2934

Table 8: Top 10 retweeted tweets

Similar to the previous category, even here we are able to find various topics mentioned with varying seriousness of how the messages of said tweets are worded. There are once again two tweets mentioning the boy band BTS (2, 5). We also encounter a message by Indonesia's president Joko Widodo (8) about the need for cooperation, his call for peace and the awareness of sustainability and economic growth. However, he chooses to compare the aforementioned problems to the fictional villain Thanos from Marvel's comic and cinematic universe which are full of magic, sci-fi and fantasy elements, and superheroes. It is perhaps an interesting take on how to communicate such problems and

how to get the discussion among the general public, which is reflected in the reactions to this tweet, in the comments and quoted retweets. Finally, we see a tweet which focuses on promoting the Trillion Trees project (not to be confused with the 1t.org mentioned in the previous subchapter, even though they often cooperate) which joins together a collective effort of BirdLife International, the Wildlife Conservation Society (WCS), and World Wildlife Fund (WWF) along with other collaborating organizations. Their goal is not as literal as their name, as they focus on general local, or regional efforts focusing on reforestation of deserted and desolated areas throughout the world.

We are once again able to find different categories, where three tweets include speeches and messages from Davos 2020 (3, 6, 9) and one includes a message from 2018 ASEAN meeting (8). Further, three tweets include speeches and messages by politicians - former US President Donald Trump (3), Indonesia's president Joko Widodo (8), and Pakistan's PM Imran Khan (9), one includes a message by climate activist Greta Thunberg (6), one includes a message by a famous physicist Stephan Hawking (10), two mention a popular music boy band BTS (2, 5), two focus on popularizing and informing about a concrete climate action related project (4, 7), and one informs about India's economic growth (1).

We are once again able to link the mentioned topics to individual SDGs and we identify the following topics being mentioned: economic growth (1, 3, 8), culture (2, 5), climate change and climate action (4, 6), just and peaceful nations (9), recycling and sustainability (4, 6), technology and science (10), and call for global cooperation and globalization (2, 3, 4, 5, 6, 8, 10).

Further, we are able to pivot certain tweets to concrete locations, or more loosely through the people who are featured. Thus, there are two tweets regarding India (1, 7), two mentioning Pakistan (4, 9) - one of which focuses on Khyber Pakhtunkhwa province (4), and two mentioning South Korea through the boy band BTS (2, 5). Further, one tweet includes a speaker from the USA (3), Sweden (6), Indonesia (8), and the UK (10).

6.3.3 Top 10 replied to tweets

The average number of replies a tweet by WEF received over the selected time period was 3. The highest number of replies was 4654 while for the 10th place the amount of

replies was 951. The average number of replies in the top 10 is 2188 replies. As with the previous category, even here we are able to see a certain overlap. However, in this case it is just one tweet (2), Donald Trump’s speech made during the 2020 Davos meeting, which is included in both of the previous categories. The tweets can be seen in Table 9, below.

Keywords	Date	Media	SDGs	Replies
Venus, climate change, sustainability	18/12/2020	Link, Image	12, 13	4654
Donald Trump, speech, Davos 2020, economy, USA	21/01/2020	Link, Video	8, 9, 17	4478
Covid, pandemic, health, travel	30/07/2020	Link, Video	3, 13, 17	2486
Rhiane Fatinikun, travel, hiking, inclusivity, diversity	06/10/2020	Link, Image	5, 10	2404
Food, technology, sustainability, plant-based meat	27/12/2020	Link, Video	2, 12, 13, 15	1900
Covid, economy, sustainability, Great Reset project, Prince Charles, Klaus Schwab, stakeholder economy	17/11/2020	Link, Image	3, 8, 9, 11, 13, 16, 17	1462
Young Global Leaders, human rights, equality, economy, society, youth	13/03/2019	Link, Image	4, 5, 8, 9	1394
Food, technology, sustainability, plant-based meat	22/11/2020	Link, Video	2, 12, 13, 15	1106
Technology, digital ID, social security	25/12/2020	Link, Image	3, 5, 8, 9	1042
Climate change, carbon emissions, Bill Gates, Gates Foundation	12/11/2020	Link, Video	12, 13, 17	951

Table 9: Top 10 replied to tweets

In this category we are once again able to see varying topics. However, in this case they are all connected by a common factor - all had raised some level of emotions, positive or negative, which made people react and voice their opinion through replies. As such, while the tweets attempt to raise serious questions about economy, social inequalities, technological progress, environment, or global health, through the replies they are often skewed and twisted, in certain cases to ideas which seem conspiratory. Once again we are able to see two tweets featuring the same topic with a little change to the content (5,

8), with the less popular including the #BoldActions4Food hashtag, but otherwise being the same.

As an example, the most replied to tweet informs and links to an article about the atmosphere on the planet Venus being Earth-like, but how climate change made it uninhabitable a long time ago. However, the WEF is a proponent of the idea that the climate change on Earth is caused mostly by humans. Thus, many of the replies mock the tweet by mentioning “Venusians” and how Venus lacked plans such as The Paris Agreement, Green Deal for Europe, or the Great Reset. They attempt to bring up WEF’s alleged hypocrisy. Similarly, we can read about the global elite trying to control the world's citizens through vaccines (3), tracking technology (9), or other means (6, 10). Some tweets feature people with controversial reputations (2, 7), or bring up sensitive topics about racial identity, racism and inclusivity (4), or about food, vegetarianism and plant-based meat (5, 8). Some of the tweets also focus, or mention the ongoing Covid-19 pandemic resulting in controversial claims about the disease, or how it rapidly transformed the way we think about travel, global interconnectedness, or economy and society in general (3, 6, 9, 10).

In these tweets we can once again encounter a speech from WEF’s annual meeting in Davos from 2020 by former US President Donald Trump (2), and the idea proposed by Prof. Schwab for Davos 2021 with the overall topic called the Great Reset (6). Further, we see a WEF aligned project Young Global Leaders (6), focusing on successful and influential youth from all over the globe from activists and politicians, to entrepreneurs and businesspersons. One of the tweets mentions the UN’s Framework Convention on Climate Change, along with the Gates Foundation and Green Horizon Summit, which preceded the 2020’s UN Climate Change Conference (10).

Two of the tweets include a political leader - former US President Donald Trump (2), and the UK’s Prince Charles (6), two include influential managers or thinkers (6, 7), and two include a climate and a social justice activism, one featuring Bill Gates and the Gates Foundation (10), and one featuring Rhiane Fatinikun (4), founder of Black Girls Hike and advocate for social inclusivity for black women. Further, five tweets focus on informing about topics related to: climate change (1), pros of digital identity (3, 9), sustainable food production (5, 8).

The connection between the content and topic of the tweets and individual SDGs was once again considered and the goals are listed in the table above. Further, pivotal topics of the tweets were identified: technology (3, 5, 8, 9), climate change (1, 10), sustainable consumption (5, 8), digital identity (3, 9), economy and economic growth (2, 6), equal rights and opportunities (4, 7), and call for global cooperation and globalization (1, 2, 3, 5, 6, 7, 8, 9, 10).

However, in the case of replied tweets most do not have any firm geographical pivot point. One tweet focuses on equal rights activism in the UK (4). Further, we are able to loosely connect two tweets to Switzerland's Davos (2, 6), one mentions Japan and Tokyo's Summer Olympics (3), two feature scientists from Israel (5, 8), two tweets to speakers from the USA (2, 10) and one to an activist from the UK (4).

6.3.4 Top 10 quote retweeted tweets

The average number of quoted retweets for the WEF's Twitter profile was 6. The average number of quoted retweets in the top 10 most quote retweeted tweets was 1891. The highest amount of quote retweets is 3225, while the 10th most quoted retweet has 974 quoted retweets. Once again we are able to identify certain topical overlap with the previous categories. To be more precise, with the category of the most replied to tweets. In this case it is the total number of 6 tweets repeated in this category as well. The tweets can be seen in Table 10, on the next page.

Once again we are able to identify a varied array of topics contained within these tweets. However, in this case we are able to identify a dominant and prevailing one among them – health, both through the lenses of technological progress and through the sustainable and alternative production of food. Further, similar to the previous category even here we are able to see replies and quotes that hint towards a stronger emotional engagement with the content of the original tweet. Even here we are able to identify tweets with very similar, or the same content. The two tweets featuring a plant-based meat (4, 6) and the two tweets considering ethics of fatal car crashes decision making for self-driving cars (2, 7).

Keywords	Date	Media	SDGs	Quotes
Venus, climate change, sustainability	18/12/2020	Link, Image	12,13	3225
Technology, ethics, health, traffic	03/11/2018	Link, Image	3, 9	2906
Covid, pandemic, health, travel	30/07/2020	Link, Video	3, 13, 17	2751
Food, technology, sustainability, plant-based meat	27/12/2020	Link, Video	2, 12, 13, 15	2658
Technology, health	08/01/2018	Link, Image	3, 9	1546
Food, technology, sustainability, plant-based meat	22/11/2020	Link, Video	2, 12, 13, 15	1444
Technology, ethics, health, traffic	27/12/2018	Link, Image	3, 9	1207
Food, sustainability, health	27/11/2020	Link, Image	3, 12	1116
Covid, economy, sustainability, Great Reset project, Prince Charles, Klaus Schwab, stakeholder economy	17/11/2020	Link, Video	3, 8, 9, 11, 13, 16, 17	1082
Technology, digital ID, social security	25/12/2020	Link, Video	3, 5, 8, 9	974

Table 10: Top 10 quote retweeted tweets

From the 4 tweets which are original in this category only 3 are unique, as 2 of those (2, 7) overlap topically with the same link and a very similar wording. Once again we are able to see extensive discussion about the implications of those tweets, or people hinting at WEF's possible hypocrisy, or misunderstanding of certain topics. As an example, the tweets regarding self-driving cars within the context of fatal accidents present a graph in the attached image. The study was based on AI decision making - if the crash was unavoidable with almost certain fatality, would the car protect its driver by hitting a person in the road, or would it protect the person in the road by, for example, aiming for a wall killing the driver. Further, those were divided into multiple categories - stroller, boy, girl, pregnant person, homeless person, old man/woman, dog, cat, criminal, female/male doctor, etc. The graph then shows the results of the hypothetical situations with how likely, compared to the average; it is for a certain group to be the victim of

such a crash. The discussion, along with the quote retweets, focuses on certain aspects of the presented results. One group misunderstands the meaning of the data or purposefully draws misleading conclusions, one group ponders certain combinations (ex. Homeless male athlete pushing a stroller), one group points out the possible implications and social acceptance of certain comparisons: athletes, who are less likely to get hit, versus large bodied people, who are more likely to get hit; similarly people in the executive positions are safer than homeless people, and dogs are safer than cats. The final group then ponders how it would be possible for the AI to recognize whether a person is a doctor, in the executive, or if that person is a criminal. Similar variety of reactions can then be seen in the other two tweets - one about a silicone heart transplant (5), where one group praises the invention, others light-heartedly talk about “unbreakable heart”, and some mention a fictional character Cristina Yang from a series called Grey’s Anatomy (FANDOM TV Community, n.d.), who allegedly invented a similar device within the series, and the other about a pie-graph of a balanced diet (8) where most of the people disagree with WEF in general, or at least with information shown in the graph, often adding their own guides, or tips for a healthy diet.

Overall, we are able to see tweets which focus on the topic of health, often connected to technology (2, 3, 5, 7, 10), or the broader topic of sustainability and consumption (4, 6, 8). Similarly, we can see tweets regarding the Covid-19 pandemic (3, 9, 10). Further, we can once again see Prof. Schwab’s idea for Davos 2021 and its topic called the Great Reset (9). We are also able to identify three tweets (4, 6, 8) which focus on a healthy diet and sustainable food and which are related to the WEF aligned summit from November 2020 called Bold Actions for Food as a Force for Good.

The respective connection between each tweet and individual SDGs are listed in the table above. Further, certain pivotal and important topics were identified: health (2 to 10), technology (2 to 7, 9, 10), sustainability (1, 4, 6, 8, 9), digital identity (2, 10), food and diet (4, 6, 8), AI (2, 7), climate change (1), and the call for global cooperation and globalization (1, 3, 4, 6, 8, 9, 10).

Similar to the previous categorization, we are able to loosely connect certain tweets to concrete geographical locations. None of the tweets explicitly mentions a concrete place. However, two tweets mention a research group from the US (2, 7), two mention

researchers from Israel (4, 6), and two can be connected to Switzerland, one to Zurich (5) and one to Davos (9).

7 Conclusions, findings and further work

This chapter serves as an overall summary for the previous chapter, establishing the main areas of interest of the WEF's activity on Twitter based on the dataset of tweets between the years 2016 and 2020.

Further, it addresses some of the issues with the data which were found during the work on this thesis. Those were in certain cases addressed and mitigated, in others considered. However, they may pose a challenge for future research using the same or similar inputs or methodology.

Lastly, we provide some suggestions for future work.

7.1 Geographical content

When assessing the popularity of regions by countries or regions, we were able to identify a strong group of four countries - China, India, United States of America and Japan. While the first three countries' popularity can be explained by them being the most populous countries in the world and thus offering more opportunities and topics to tweet about, in the case of Japan the share of Twitter users comes into focus. According to statistics, Japan has the second largest number of users just after the United States with many individuals and companies using the platform for personal and business purposes. Thus, perhaps Japan is an easy target for online marketing by the WEF. Further, all of those countries are global players in politics, innovation and technology.

Other interesting results come from the countries of northern Europe. Compared to the rest of the countries on top these have a relatively low population. These countries are - Sweden, Finland, Norway, Denmark, and Iceland. The topics which are often mentioned in tweets within the context of those countries are focusing strongly on the social and environmental aspects. Those include: high population happiness, sustainability, recycling and clean energy, well-being and quality work environment, social stability and cohesion, and social inclusivity and gender equality. Further, in the case of Finland the topic of education is a significant factor, while for Norway it is their approach to prison management, and for Iceland it is the unique possibility to utilize geothermal energy and potential for tourism.

Similarly, while examining the dataset by continents and the number of mentions, albeit by a simple and naive approach when regarding the continent-related keywords, we are able to identify an important difference. When we summarize the mentions by countries, Africa amounts to 4.72% of all mentions, followed by only Oceania and Antarctica. However, when we use the continent names instead, Africa is second with 27.18% of the mentions. It is perhaps the case, where the whole region of Africa is often treated as one, instead of mentioning the problems of countries individually. On the contrary, the continent of Asia by keywords amounts to 7.36% of the mentions, while by countries it covers the 49.12% of the mentions - with China, India, and Japan being among the most popular countries. Further implications of these results can be drawn after the proper re-assessment of the continent-related keywords and the examination of the overall context of those tweets.

7.2 Popular and important topics

In this case we measured the total number of tweets in which certain hashtags, as determinants of a topic, and users were mentioned.

This way we were able to identify the most popular categories of topics the WEF tweets about most often. Among those we were able to identify three hashtags which addresses the current ongoing global thread of the COVID-19 pandemic - #health, #covid19, and #coronavirus. Further, we identified concrete regions - #china, #europe, #us, #africa, and #india. The popularity of the #africa among others related to concrete countries further supports the idea of the region being treated as one. The WEF tweets concise as well of lists in various categories labeled with the hashtag #bestof. From the perspective of popularity the hashtag #wef17 related to the 2017 meeting in Davos is the most popular, followed by similar hashtags for years 2019 and 2018. However, it is perhaps the change of tweeting pattern by the WEF that has led to this result.

However, the vast majority of other popular hashtags are related to important topics of sustainability and sustainable development. We identified 4 main categories. Three of those follow the pillars of sustainable development – societal, environmental and economic, while the fourth has emerged naturally as a connecting factor for the previous three – technology. We attempted to further connect those topics to the individual SDGs. However, it is important to note that those are highly intertwined and

as such those may overlap into other categories as well. We decided not to include the Goal #17: Partnership for the goals, as it is in essence contained in the vast majority of the WEF's tweets on various topics. Based on the topics we identified that the most often mentioned SDGs were the #9: Industry, innovation and infrastructure, #12: Responsible consumption and production, and #8: Decent work and economic growth. On the contrary, the least popular topics were #1: No poverty, #6: Clean water and sanitation, #16: Peace, justice and strong institutions, and #2: Zero Hunger.

When examining the most often mentioned users we were able to identify multiple types based on their connection to the WEF. Some mentions are of the accounts related to the WEF – secondary accounts related to concrete topic, or people directly connected to the WEF as employees. Further, we were able to find influential organizations and individuals who often share similar view, or talk about similar topics as the WEF. The third group was the profiles connected to companies and organizations from which the WEF takes the data for their visualizations.

7.3 Most popular tweets

When we focused on the number of reactions for tweets we were able to identify the most popular tweets for each type of interaction. For this purpose we chose the top 10 most popular tweets in each category.

We were able to connect those tweets to concrete topics, SDGs and in some cases to concrete geographical locations or people. Further, we were able to identify certain topical trends for some of the categories. Namely, when considering the most popular tweets by likes and by retweets some of those were included in both categories, or were quite alike topically. Similarly, for the other two categories of replies and quoted retweets we observed the same thing.

If we were to consider the overall sentiment of each tweet in those two categories, we could say that in the case of the first one the overall reactions and contents to those tweets were positive, while in the case of the second group the tweets often contained topics and ideas which were deemed controversial, or wrong by the people who decided to react. It is also important to note that both like and retweet require less user input

than replying and retweeting with a quote – as those include their own input into the problematic of each topic.

7.4 Overlapping meanings of keywords

Keywords were often a pivotal element for the content analysis of each tweet. It was often not one keyword related to one term, but rather a group of keywords which covered the desired topic accurately enough. When we wanted to count all tweets which were related to China, we searched for the other forms of the word as well, sometimes with other considerations, such as the name of the population living there.

However, this approach comes with a caveat – the word, or abbreviation used may not have just one meaning. For example, using the “us” as a keyword for the tweets related to the USA we get the results which use the desired meaning, but the other meaning of the word as well: word used to refer to a speaker and one or more other people. While we were able to address this issue when considering countries, this pose a challenge to future analysis of the tweets; especially if we were to focus on the words which commonly have multiple meanings based on the context.

7.5 Repetitiveness of tweets

Further, delving deeper into certain subgroups of tweets we were able to identify certain patterns in WEF’s posts. Averaging the almost 71 posts per day, it is perhaps no surprise that tweets regarding certain topics and regions were often reposted multiple times, sometimes even within days from the posting of the original tweet. AS an example, in the category of the 10 most liked tweets there are 3 similar tweets about the same topics which link to the same article on the WEF’s website. While this approach hints at a certain emphasis on those topics by the WEF, it might skew the outlook of overall popularity of certain topics as the amounts of interactions with those tweets are dispersed across many similar or exactly the same tweets. Similarly, the information about the use of certain keywords, hashtags, or user mentions is affected as well.

As such, we suggest further examination of these post patterns, where the overall similarity between the tweets should be assessed. However, this approach would require a development of sophisticated methodology, as the mere comparison for equality will not suffice. In many cases the texts of tweets in question are not exactly the same - they

differ in the use of hashtags, or the exact web address of a link; even though those often redirect to the same article. For such a purpose the utilization of string metrics, such as the Levenshtein distance, is advised.

To simply demonstrate this possible issue we can use Twitter's Search functionality with the query *until:2020-12-31 since:2016-01-01 from:wef "taiwan"*. Sorting the results by Latest there are 19 tweets - we can argue, however, that only two of those are original. One regards the banning of plastic by 2030 (WEF, 2018), while the other eighteen all mention the same thing. Those tweets focus on a newly built "twisting tower block with nearly as many plants as Central Park" (WEF, 2017). All of these tweets include a link to the same article at the WEF's own website. However, the link is different each time. There are also two versions of the rest of the text - one version contains a typo, while the other has the word corrected.

7.6 Suggestions for further work

The amount of data we were able to obtain through the utilization of web scraping and further data processing is quite vast. As such, we were able to analyze and process but a fraction of what may be possible in the future.

The scope through which the data are examined is also quite important. While in this case we focused on the topics of sustainability and geography, the data may as well be used for a wider range of topics. Some examples of relevant topics are sentiment analysis of tweets and reactions, the connection between the data in tweets and the WEF's website, or deeper lexical analysis of the text. Further, the tool which we used for the acquiring of our dataset may be utilized for other purposes as well – analysis of other profiles on other social media through similar approach.

Finally, we hope that this thesis proved that the data contained and available on the social media can be utilized as an input for the academic work across different fields and not just the informatics. We hope to focus on the topics similar to this in the future as well.

Conclusion

Due to WEF's prolific activities on Twitter we were able to acquire the dataset of almost 130 thousands tweets from 2016 to 2020 which were used as an input for further analysis. This dataset includes information about each individual tweet – URL address, content, date and time, number of interactions, lists of included links and media, hashtags and user mentions, and information about place and coordinates.

We used this dataset as an input for our further work. To simplify the work with a dataset of this size we custom developed a program in the programming language Java to help with the data processing.

First, we derived keywords related to the names of countries and continents to present the list of 200 countries and 6 continents and their popularity in the tweets. We provided further information about other possible approaches to similar kind of analysis.

Second, we focused on the analysis of overall topics of tweets using hashtags and mentioned users. We present a list of 27 most popular hashtags used by the WEF and a list of 20 most often mentioned Twitter users. We further found the ties between the topics and idea of sustainability and the SDGs. We also propose other interesting or relevant hashtags and users.

Third, we analyze the tweets by their popularity measured by the amount of interactions they received – likes, retweets, replies and quoted retweets. We were able to identify two main groups of tweets – likes and retweets with overall positive sentiments, and replies and quoted retweets with overall controversial or negative leaning sentiments of reactions. Further, we identified users, topics and SDGs related to each individual tweet.

The definition of reliable keywords proved to be a challenge, as certain desired words and phrases may be presented in a different context and meaning than what we are looking for. We came across this issue when assessing the data from geographical point of view. While we were able to reliably control the results for country names, we had to step down from the keyword analysis of individual SDGs. Those and similar analysis focusing on words and phrases with more fuzzy meaning would require a development of additional methodological framework to address the context of the results and other issues.

We conclude that utilization of publicly, or semi-publicly available data for this kind of analysis and work is possible and may yield results for multiple purposes. From the here presented geographical information and topical information to the possible outreach to the fields of marketing, sentiment analysis, or some deeper analysis utilizing artificial intelligence and machine learning technologies; some of which are also available publicly.

List of the Bibliographic References

- Ballestar, M. T., Cuerdo-Mir, M., & Freire-Rubio, M. T. (2020). The concept of sustainability on social media: A social listening approach. *Sustainability*, 12(5), 2122.
- Ban, K. M. (2019). *Sustainable Development Goals*. GRIPS Forum 2018. https://www.grips.ac.jp/cms/wp-content/uploads/2018/12/190314_penseeSP_en.pdf
- Calcagni, F., Amorim Maia, A. T., Connolly, J. J. T., & Langemeyer, J. (2019). Digital co-construction of relational values: Understanding the role of social media for sustainability. *Sustainability Science*, 14(5), 1309-1321.
- Caradonna, J. L. (2014). *Sustainability: A history*. Oxford University Press.
- Carlowitz, H. C. E. von (1713): *Sylvicultura Oeconomica*. Meissen.
- CERN. (n.d.). <http://info.cern.ch> - home of the first website. <http://info.cern.ch>
- Coe, N. M., Kelly, P. F., & Yeung, H. W. C. (2013). *Economic Geography: A Contemporary Introduction* (2nd ed.). Wiley.
- Cohen, M. (2017). A systematic review of urban sustainability assessment literature. *Sustainability*, 9(11), 2048.
- Dicken, P. (2015). *Global Shift: Mapping the Changing Contours of the World Economy* (Seventh ed.). SAGE Publications Ltd.
- Eclipse Foundation. (n.d.) *Enabling Open Innovation & Collaboration*. <https://www.eclipse.org/>
- Ehrenfeld, J. R. (2008). Sustainability by design. *In Sustainability by Design*. Yale University Press.
- Elkington, J. (2006). Governance for sustainability. *Corporate governance: an international review*, 14(6), 522-529.
- Elsevier. (2022a). *Journal of Cleaner Production – Journal – Elsevier*. Retrieved March 20, 2022, from <https://www.journals.elsevier.com/journal-of-cleaner-production>
- Elsevier. (2022b). *Journal of Environmental Management – Journal – Elsevier*. Retrieved March 20, 2022, from <https://www.journals.elsevier.com/journal-of-environmental-management>
- ESRI. (n.d.) *ArcGIS Desktop | Desktop GIS Software Suite – esri*. Retrieved March 22, 2022, from <https://www.esri.com/en-us/arcgis/products/arcgis-desktop/overview>
- Faber, N., Jorna, R., & Van Engelen, J. O. (2010). The sustainability of “sustainability”—A study into the conceptual foundations of the notion of “sustainability”. *In Tools, techniques and approaches for sustainability: Collected writings in environmental assessment policy and management* (pp. 337-369).
- FANDOM TV Community. (n.d.). *Christina Yang | Grey’s Anatomy Universe Wiki | Fandom*. Retrieved April 10, 2022, from https://greysanatomy.fandom.com/wiki/Cristina_Yang

- Force, A., Manuel-Navarrete, D., & Benessaiah, K. (2018). Tourism and transitions toward sustainability: developing tourists' pro-sustainability agency. *Sustainability Science*, 13(2), 431-445.
- Freeman, E., & Liedtka, J. (1997). Stakeholder capitalism and the value chain. *European Management Journal*, 15(3), 286-296.
- GG Gap (2017). *World economic forum*. Cologne/Geneva.
- Goddard, N. D. R., Kemp, R. M. J., & Lane, R. (1997). An overview of smart technology. *Packaging Technology and Science: An International Journal*, 10(3), 129-143.
- Google. (n.d.). *Google Sheets: Free Online Spreadsheet Editor*. Retrieved April 7, 2022, from <https://www.google.com/sheets/about/>
- Goodland, R. (1995). The concept of environmental sustainability. *Annual review of ecology and systematics*, 26(1), 1-24.
- Graz, J. C. (2003). How powerful are transnational elite clubs? The social myth of the World Economic Forum. *New Political Economy*, 8(3), 321-340.
- Grebenyuk, Y. (2022). *Will Smith and Chris Rock's 2022 Oscars Incident Over Jada Pinkett Smith: Everything to Know*. US Magazine. Retrieved March 31, 2022, from <https://www.usmagazine.com/entertainment/pictures/will-smith-and-chris-rocks-oscars-2022-incident-everything-to-know/the-exchange/>
- Hák, T., Janoušková, S., & Moldan, B. (2016). Sustainable Development Goals: A need for relevant indicators. *Ecological indicators*, 60, 565-573.
- Hamid, S., Ijab, M. T., Sulaiman, H., Anwar, R. M., & Norman, A. A. (2017). Social media for environmental sustainability awareness in higher education. *International Journal of Sustainability in Higher Education*.
- Ilieva, R. T., & McPhearson, T. (2018). Social-media data for urban sustainability. *Nature Sustainability*, 1(10), 553-565.
- International Telecommunication Union. (2019). *Measuring digital development: Facts and Figures 2019*. ITUPublications. <https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2019.pdf>
- International Telecommunication Union. (2021). *Measuring digital development: Facts and Figures 2021*. ITUPublications. <https://www.itu.int/en/myitu/Publications/2021/11/25/14/45/Facts-and-figures-2021>
- International Telecommunication Union. (n.d.). *ITU: Committed to connecting the world*. Retrieved March 11, 2022, from <https://www.itu.int/en/Pages/default.aspx>
- Introducing JSON. (n.d.). <https://www.json.org/json-en.html>
- JustAnotherArchivist. (n.d.). *snsrape* [GitHub Repository]. <https://github.com/JustAnotherArchivist/snsrape>
- Kaplan, A. M., & Haenlein, M. (2009). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59-68. <https://doi.org/10.1016/j.bushor.2009.09.003>

- Keiner, M. (2005). *History, definition(s) and models of sustainable development*. ETH Zürich. <https://doi.org/10.3929/ethz-a-004995678>
- Le Blanc, D. (2015). Towards integration at last? The sustainable development goals as a network of targets. *Sustainable Development*, 23(3), 176-187.
- Liverman, D. M. (2018). Geographic perspectives on development goals: Constructive engagements and critical perspectives on the MDGs and the SDGs. *Dialogues in Human geography*, 8(2), 168-185. <https://doi.org/10.1177/2043820618780787>
- Manetti, G., & Bellucci, M. (2016). The use of social media for engaging stakeholders in sustainability reporting. *Accounting, Auditing & Accountability Journal*.
- Marcuse, P. (1998). Sustainability is not enough. *Environment and urbanization*, 10(2), 103-112.
- Mazanec, J. A., & Ring, A. (2011). Tourism destination competitiveness: second thoughts on the World Economic Forum reports. *Tourism Economics*, 17(4), 725-751.
- Mensah, J., & Casadevall, S. R. (Reviewing editor) (2019). Sustainable development: Meaning, history, principles, pillars, and implications for human action: Literature review. *Cogent Social Sciences*, 5(1). <https://doi.org/10.1080/23311886.2019.1653531>
- Natural Earth. (2022). *Natural Earth – 1:10m Cultural Vectors*. Retrieved April 20, 2022, from <https://www.naturalearthdata.com/downloads/10m-cultural-vectors/>
- Nilsson, M., Griggs, D., & Visbeck, M. (2016). Policy: map the interactions between Sustainable Development Goals. *Nature*, 534(7607), 320-322.
- OECD. (2007). *Participative web and user-created content web 2.0, Wikis and social networking*. OECD Publishing. <https://doi.org/10.1787/9789264037472-en>
- Oracle. (n.d.). *What is Java technology and why do I need it?* Retrieved November 11, 2021, from https://www.java.com/en/download/help/whatis_java.html
- Pigman, G. A. (2007). *The World Economic Forum: A multi-stakeholder approach to global governance*. Routledge.
- Portney, K. E. (2015). *Sustainability*. MIT Press.
- Reilly, A. H., & Hynan, K. A. (2014). Corporate communication, sustainability, and social media: It's not easy (really) being green. *Business horizons*, 57(6), 747-758.
- Rodríguez-Olalla, A., & Avilés-Palacios, C. (2017). Integrating sustainability in organisations: an activity-based sustainability model. *Sustainability*, 9(6), 1072.
- Sachs, J. D. (2012). From millennium development goals to sustainable development goals. *The lancet*, 379(9832), 2206-2211.
- Schwab, K. (1973). *Davos manifesto 1973: A code of ethics for business leaders* [Online Version]. World Economic Forum. Retrieved March 14, 2022, from <https://www.weforum.org/agenda/2019/12/davos-manifesto-1973-a-code-of-ethics-for-business-leaders/>
- Schwab, K. (2020). *Davos manifesto 2020: The universal purpose of a company in the Fourth Industrial Revolution* [Online Version]. World Economic Forum. Retrieved March 14, 2022, from <https://www.weforum.org/agenda/2019/12/davos-manifesto-2020-the-universal-purpose-of-a-company-in-the-fourth-industrial-revolution>

- Schwab, K. (2021). *Stakeholder capitalism: A global economy that works for progress, people and planet*. John Wiley & Sons.
- Scoones, I. (2007). Sustainability. *Development in practice*, 17(4-5), 589-596.
- Shafranovich, Y. (2005). *Common Format and MIME Type for Comma-Separated Values (CSV) Files* [Memo]. <https://datatracker.ietf.org/doc/html/rfc4180>
- Social Blade. (n.d.a). *World Economic Forum's Twitter Stats Summary Profile (Social Blade)*. Retrived March 12, 2022, from <https://socialblade.com/twitter/user/wef>
- Social Blade. (n.d.b). *All About Social Blade*. Retrived March 12, 2022, from <https://socialblade.com/info>
- Social Blade. (n.d.c). *Top 50 Twitter users sorted by Followers*. Retrived March 12, 2022, from <https://socialblade.com/twitter/top/50/most-followers>
- Springer, M., & Schnelzer, J. (2019). *Differentiation of Industry 4.0 Models: The 4th Industrial Revolution from different Regional Perspectives in the Global North and Global South*. <https://dx.doi.org/10.13140/RG.2.2.35510.55363>
- Statista. (2022a). *Most used social media 2021 | Statista*. Retrieved April 22, 2022, from <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
- Statista. (2022b). *Twitter: most users by country | Statista*. Retrieved April 22, 2022, from <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>
- Stevens, T. M., Aarts, N., Termeer, C. J. A. M., & Dewulf, A. (2016). Social media as a new playing field for the governance of agro-food sustainability. *Current Opinion in Environmental Sustainability*, 18, 99-106.
- Sustainability. (2022). *Sustainability | An Open Access Journal from MDPI*. <https://www.mdpi.com/journal/sustainability>
- Sustainability Science. (2022). *Sustainability Science | Home*. <https://www.springer.com/journal/11625>
- Thiele, L. P. (2016). *Sustainability*. John Wiley & Sons.
- Twitter. (n.d.a). Home / Twitter [Twitter Homepage]. Twitter. Retrieved April 22, 2022, from <https://twitter.com/home>
- Twitter. (n.d.b). *Twitter API Documentation | Docs | Twitter Developer Platform*. Developer Twitter. Retrieved April 22, 2022 from <https://developer.twitter.com/en/docs/twitter-api>
- United Nations. (2015a). *The Millenium Development Goals Report 2015*. [https://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20\(July%201\).pdf](https://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf)
- United Nations. (2015b). *Transforming Our World: The 2030 Agenda for Sustainable development*. General Assembly, 70th session. Retrieved March 16, 2022, from <https://sdgs.un.org/2030agenda>
- United Nations. (n.d.a). *About us*. United Nations. Retrieved February 13, 2022, from <https://www.un.org/en/about-us>

United Nations. (n.d.b) *Sustainable Development*. United Nations. Retrieved March 17, 2022, from <https://sdgs.un.org/>

United Nations Development Programme. (2010). *What Will It Take to Achieve the Millennium Development Goals? - An International Assessment*. <https://www.oecd.org/dac/gender-development/45604012.pdf>

United Nations General Assembly. (2000). *United Nations Millennium Declaration*. New York. <http://www.un.org/millennium/declaration/ares552e.htm>

United Nations Statistics Division. (2021). *UNSD — Methodology*. UNSD. Retrieved March 17, 2022, from <https://unstats.un.org/unsd/methodology/m49/>

United Nations Statistics Division. (n.d.). *SDG Indicators – SDG Indicators*. Retrieved March 18, 2022, from <https://unstats.un.org/sdgs/indicators/indicators-list/>

Ward, I. (n.d.) *JSON Lines: Documentation for the JSON Lines text file format*. <https://jsonlines.org/>

Waughray, D. (2011). *Water securitythe water-food-energy-climate nexus: the World Economic Forum water initiative* (No. 333.91 W3).

WEF Cybersecurity [@WEFCybersec]. (n.d.). *Tweets* [Twitter profile]. Twitter. Retrieved March 12, 2022, from <https://twitter.com/WEFCybersec>.

WEF Energy [@wefenergy]. (n.d.). *Tweets* [Twitter profile]. Twitter. Retrieved March 12, 2022, from <https://twitter.com/wefenergy>

Wilkinson, A., Hill, M., & Gollan, P. (2001). The sustainability debate. *International Journal of Operations & Production Management*.

World Economic Forum [@wef]. (2017, August 27). *This twisting Taiwan tower block has nearly has many plants as Central Park* <http://wef.ch/2g46Eqw> [Image attached] [Tweet]. Twitter. <https://twitter.com/wef/status/901914478530494465>

World Economic Forum [@wef]. (2018, February 27). *#Taiwan has committed to banning plastic items by 2030* <http://wef.ch/2sYKps9> #environment [Image attached] [Tweet]. Twitter. <https://twitter.com/wef/status/968455951568261120>

World Economic Forum [@wef]. (2020, January 21). *US President @realDonaldTrump spoke at #wef20, telling the audience the United States is in the midst of an economic boom*. [Video attached] [Tweet]. Twitter. <https://twitter.com/wef/status/1219594713470263296>

World Economic Forum. (n.d.a). *The World Economic Forum*. Retrieved March 14, 2022, from <https://www.weforum.org/>

World Economic Forum. (n.d.b). *Reports | The World Economic Forum*. Retrieved April 20, 2022, from <https://www.weforum.org/reports/>

World Economic Forum [@wef]. (n.d.c) *Tweets* [Twitter profile]. Twitter. Retrieved March 12, 2022, from <https://twitter.com/wef>

World Economic Forum [@Davos]. (n.d.d) *Tweets* [Twitter profile]. Retrieved March 12, 2022, from <https://twitter.com/Davos>

World Economic Forum. (n.d.e). *Timeline – World Economic Forum* [Interactive website]. Retrieved April 22, 2022, from <https://timeline.weforum.org/>.

Xu, M., David, J. M., & Kim, S. H. (2018). The fourth industrial revolution: Opportunities and challenges. *International journal of financial research*, 9(2), 90-95.

Zenya, A., & Nystad, Ø. (2018). Assessing Corporate Sustainability with the Enterprise Sustainability Evaluation Tool (E-SET). *Sustainability*, 10(12), 4661.

List of Tables

Table 1: Continent mentions as sums by countries	58
Table 2: Continent mentions by keywords	60
Table 3: Sub-subregion mentions by countries.....	62
Table 4: A list of the most popular hashtags	64
Table 5: SDGs and related hashtags.....	66
Table 6: A list of the most popular mentions.....	67
Table 7: Top 10 liked tweets.....	71
Table 8: Top 10 retweeted tweets	73
Table 9: Top 10 replied to tweets.....	75
Table 10: Top 10 quote retweeted tweets.....	78

List of Figures

Figure 1: Example of a Tweet (WEF, 2020)	27
Figure 2: Graphical interface of Eclipse IDE	34
Figure 3: The WEF website, Featured topics (WEF, n.d.a)	41
Figure 4 Example of a website's visual representation and its source code	48
Figure 5: Number of tweets per month.....	51
Figure 6: Daily posts by years, 7 days moving average	52
Figure 7: Country occurrences in tweets, % of total (ArcMap).....	55
Figure 8: M49 UNSD Sub-subregions (Wikipedia, n.d.).....	61

List of Abbreviations and Symbols

4IR – 4th Industrial Revolution

API – Application Interface

CSV – Comma-separated values

IDE – Integrated Development Environment

IR – Industrial Revolution

JSON – JavaScript Object Notation

JSONL – JavaScript Object Notation Lines

MDGs – Millennium Development Goals

OECD – Organisation for Economic Co-operation and Development

SDGs – Sustainable Development Goals

UCC – User Created Content

UGC – User Generated Content

UK – United Kingdoms

UN – United Nations

UNDP – United Nations Development Programme

UNGA – United Nations General Assembly

USA – United States of America

US – United States

WBG – World Bank Group

WEF – World Economic Forum

WHO – World Health Organization

List of Appendices

Appendix A: Full table with country data

Appendix A: Full table with country data

Country / Region name	Mentions	% of total	Continental region (M49)	Sub-subregion (M49)
<i>Afghanistan</i>	25	0,13%	Asia	Southern Asia
<i>Albania</i>	0	0,00%	Europe	Southern Europe
<i>Algeria</i>	4	0,02%	Africa	Northern Africa
<i>Andorra</i>	6	0,03%	Europe	Southern Europe
<i>Angola</i>	8	0,04%	Africa	Middle Africa
<i>Antarctica</i>	118	0,64%	Antarctica	Antarctica
<i>Antigua and Barbuda</i>	0	0,00%	Americas	Caribbean
<i>Argentina</i>	45	0,24%	Americas	South America
<i>Armenia</i>	4	0,02%	Asia	Western Asia
<i>Australia</i>	337	1,82%	Oceania	Australia and New Zealand
<i>Austria</i>	11	0,06%	Europe	Western Europe
<i>Azerbaijan</i>	0	0,00%	Asia	Western Asia
<i>Bahrain</i>	4	0,02%	Asia	Western Asia
<i>Bangladesh</i>	126	0,68%	Asia	Southern Asia
<i>Barbados</i>	0	0,00%	Americas	Caribbean
<i>Belarus</i>	0	0,00%	Europe	Eastern Europe
<i>Belgium</i>	34	0,18%	Europe	Western Europe
<i>Belize</i>	1	0,01%	Americas	Central America
<i>Benin</i>	0	0,00%	Africa	Western Africa
<i>Bhutan</i>	25	0,13%	Asia	Southern Asia
<i>Bolivia</i>	3	0,02%	Americas	South America
<i>Bosnia and Herzegovina</i>	0	0,00%	Europe	Southern Europe
<i>Botswana</i>	2	0,01%	Africa	Southern Africa
<i>Brazil</i>	97	0,52%	Americas	South America
<i>Brunei</i>	11	0,06%	Asia	South-Eastern Asia
<i>Bulgaria</i>	0	0,00%	Europe	Eastern Europe
<i>Burkina Faso</i>	0	0,00%	Africa	Western Africa
<i>Burundi</i>	2	0,01%	Africa	Eastern Africa
<i>Cabo Verde</i>	0	0,00%	Africa	Western Africa
<i>Cambodia</i>	23	0,12%	Asia	South-Eastern Asia
<i>Cameroon</i>	0	0,00%	Africa	Middle Africa
<i>Canada</i>	284	1,53%	Americas	Northern America

<i>Central African Republic</i>	0	0,00%	Africa	Middle Africa
<i>Colombia</i>	58	0,31%	Americas	South America
<i>Comoros</i>	0	0,00%	Africa	Eastern Africa
<i>Costa Rica</i>	55	0,30%	Americas	Central America
<i>Croatia</i>	22	0,12%	Europe	Southern Europe
<i>Cuba</i>	38	0,20%	Americas	Caribbean
<i>Cyprus</i>	6	0,03%	Asia	Western Asia
<i>Czechia</i>	40	0,22%	Europe	Eastern Europe
<i>Democratic Republic of the Congo</i>	0	0,00%	Africa	Middle Africa
<i>Denmark</i>	266	1,43%	Europe	Northern Europe
<i>Djibouti</i>	0	0,00%	Africa	Eastern Africa
<i>Dominica</i>	0	0,00%	Americas	Caribbean
<i>Dominican Republic</i>	8	0,04%	Americas	Caribbean
<i>East Timor</i>	0	0,00%	Asia	South-Eastern Asia
<i>Ecuador</i>	14	0,08%	Americas	South America
<i>Egypt</i>	48	0,26%	Africa	Northern Africa
<i>El Salvador</i>	0	0,00%	Americas	Central America
<i>Equatorial Guinea</i>	0	0,00%	Africa	Middle Africa
<i>Eritrea</i>	12	0,06%	Africa	Eastern Africa
<i>Estonia</i>	65	0,35%	Europe	Northern Europe
<i>Eswatini</i>	0	0,00%	Africa	Southern Africa
<i>Ethiopia</i>	50	0,27%	Africa	Eastern Africa
<i>Federated States of Micronesia</i>	0	0,00%	Oceania	Micronesia
<i>Fiji</i>	1	0,01%	Oceania	Melanesia
<i>Finland</i>	417	2,25%	Europe	Northern Europe
<i>France</i>	439	2,37%	Europe	Western Europe
<i>Gabon</i>	0	0,00%	Africa	Middle Africa
<i>Gambia</i>	6	0,03%	Africa	Western Africa
<i>Georgia</i>	4	0,02%	Asia	Western Asia
<i>Germany</i>	463	2,50%	Europe	Western Europe
<i>Ghana</i>	29	0,16%	Africa	Western Africa
<i>Greece</i>	87	0,47%	Europe	Southern Europe
<i>Grenada</i>	0	0,00%	Americas	Caribbean
<i>Guatemala</i>	8	0,04%	Americas	Central America
<i>Guinea</i>	1	0,01%	Africa	Western Africa

<i>Guinea-Bissau</i>	0	0,00%	Africa	Western Africa
<i>Guyana</i>	0	0,00%	Americas	South America
<i>Haiti</i>	2	0,01%	Americas	Caribbean
<i>Honduras</i>	0	0,00%	Americas	Central America
<i>Hungary</i>	3	0,02%	Europe	Eastern Europe
<i>Chad</i>	4	0,02%	Africa	Middle Africa
<i>Chile</i>	53	0,29%	Americas	South America
<i>China</i>	3430	18,50%	Asia	Eastern Asia
<i>Iceland</i>	180	0,97%	Europe	Northern Europe
<i>India</i>	2466	13,30%	Asia	Southern Asia
<i>Indonesia</i>	105	0,57%	Asia	South-Eastern Asia
<i>Iran</i>	97	0,52%	Asia	Southern Asia
<i>Iraq</i>	47	0,25%	Asia	Western Asia
<i>Ireland</i>	63	0,34%	Europe	Northern Europe
<i>Israel</i>	42	0,23%	Asia	Western Asia
<i>Italy</i>	292	1,57%	Europe	Southern Europe
<i>Ivory Coast</i>	0	0,00%	Africa	Western Africa
<i>Jamaica</i>	12	0,06%	Americas	Caribbean
<i>Japan</i>	1307	7,05%	Asia	Eastern Asia
<i>Jordan</i>	17	0,09%	Asia	Western Asia
<i>Kashmir</i>	0	0,00%	Asia	Southern Asia
<i>Kazakhstan</i>	0	0,00%	Asia	Central Asia
<i>Kenya</i>	152	0,82%	Africa	Eastern Africa
<i>Kiribati</i>	1	0,01%	Oceania	Micronesia
<i>Kosovo</i>	0	0,00%	Europe	Southern Europe
<i>Kuwait</i>	1	0,01%	Asia	Western Asia
<i>Kyrgyzstan</i>	8	0,04%	Asia	Central Asia
<i>Laos</i>	0	0,00%	Asia	South-Eastern Asia
<i>Latvia</i>	4	0,02%	Europe	Northern Europe
<i>Lebanon</i>	21	0,11%	Asia	Western Asia
<i>Lesotho</i>	0	0,00%	Africa	Southern Africa
<i>Liberia</i>	2	0,01%	Africa	Western Africa
<i>Libya</i>	1	0,01%	Africa	Northern Africa
<i>Liechtenstein</i>	0	0,00%	Europe	Western Europe
<i>Lithuania</i>	4	0,02%	Europe	Northern Europe
<i>Luxembourg</i>	1	0,01%	Europe	Western Europe

<i>Madagascar</i>	10	0,05%	Africa	Eastern Africa
<i>Malawi</i>	5	0,03%	Africa	Eastern Africa
<i>Malaysia</i>	58	0,31%	Asia	South-Eastern Asia
<i>Maldives</i>	0	0,00%	Asia	Southern Asia
<i>Mali</i>	7	0,04%	Africa	Western Africa
<i>Malta</i>	1	0,01%	Europe	Southern Europe
<i>Marshall Islands</i>	1	0,01%	Oceania	Micronesia
<i>Mauritania</i>	3	0,02%	Africa	Western Africa
<i>Mauritius</i>	0	0,00%	Africa	Eastern Africa
<i>Mexico</i>	117	0,63%	Americas	Central America
<i>Moldova</i>	0	0,00%	Europe	Eastern Europe
<i>Monaco</i>	1	0,01%	Europe	Western Europe
<i>Mongolia</i>	9	0,05%	Asia	Eastern Asia
<i>Montenegro</i>	1	0,01%	Europe	Southern Europe
<i>Morocco</i>	25	0,13%	Africa	Northern Africa
<i>Mozambique</i>	2	0,01%	Africa	Eastern Africa
<i>Myanmar</i>	24	0,13%	Asia	South-Eastern Asia
<i>Namibia</i>	6	0,03%	Africa	Southern Africa
<i>Nauru</i>	1	0,01%	Oceania	Micronesia
<i>Nepal</i>	5	0,03%	Asia	Southern Asia
<i>Netherlands</i>	351	1,89%	Europe	Western Europe
<i>New Zealand</i>	172	0,93%	Oceania	Australia and New Zealand
<i>Nicaragua</i>	8	0,04%	Americas	Central America
<i>Niger</i>	0	0,00%	Africa	Western Africa
<i>Nigeria</i>	56	0,30%	Africa	Western Africa
<i>North Korea</i>	164	0,88%	Asia	Eastern Asia
<i>North Macedonia</i>	0	0,00%	Europe	Southern Europe
<i>Northern Cyprus</i>	0	0,00%	Asia	Western Asia
<i>Norway</i>	303	1,63%	Europe	Northern Europe
<i>Oman</i>	3	0,02%	Asia	Western Asia
<i>Pakistan</i>	142	0,77%	Asia	Southern Asia
<i>Palau</i>	1	0,01%	Oceania	Micronesia
<i>Panama</i>	16	0,09%	Americas	Central America
<i>Papua New Guinea</i>	2	0,01%	Oceania	Melanesia
<i>Paraguay</i>	4	0,02%	Americas	South America

<i>Peru</i>	36	0,19%	Americas	South America
<i>Philippines</i>	25	0,13%	Asia	South-Eastern Asia
<i>Poland</i>	28	0,15%	Europe	Eastern Europe
<i>Portugal</i>	41	0,22%	Europe	Southern Europe
<i>Qatar</i>	3	0,02%	Asia	Western Asia
<i>Republic of Serbia</i>	1	0,01%	Europe	Southern Europe
<i>Republic of the Congo</i>	7	0,04%	Africa	Middle Africa
<i>Romania</i>	0	0,00%	Europe	Eastern Europe
<i>Russia</i>	85	0,46%	Europe	Eastern Europe
<i>Rwanda</i>	85	0,46%	Africa	Eastern Africa
<i>Saint Kitts and Nevis</i>	0	0,00%	Americas	Caribbean
<i>Saint Lucia</i>	11	0,06%	Americas	Caribbean
<i>Saint Vincent and the Grenadines</i>	0	0,00%	Americas	Caribbean
<i>Samoa</i>	0	0,00%	Oceania	Polynesia
<i>San Marino</i>	0	0,00%	Europe	Southern Europe
<i>São Tomé and Príncipe</i>	0	0,00%	Africa	Middle Africa
<i>Saudi Arabia</i>	130	0,70%	Asia	Western Asia
<i>Senegal</i>	14	0,08%	Africa	Western Africa
<i>Seychelles</i>	0	0,00%	Africa	Eastern Africa
<i>Sierra Leone</i>	7	0,04%	Africa	Western Africa
<i>Singapore</i>	143	0,77%	Asia	South-Eastern Asia
<i>Slovakia</i>	6	0,03%	Europe	Eastern Europe
<i>Slovenia</i>	0	0,00%	Europe	Southern Europe
<i>Solomon Islands</i>	0	0,00%	Oceania	Melanesia
<i>Somalia</i>	11	0,06%	Africa	Eastern Africa
<i>Somaliland</i>	0	0,00%	Africa	Eastern Africa
<i>South Africa</i>	99	0,53%	Africa	Southern Africa
<i>South Korea</i>	274	1,48%	Asia	Eastern Asia
<i>South Sudan</i>	28	0,15%	Africa	Eastern Africa
<i>Spain</i>	129	0,70%	Europe	Southern Europe
<i>Sri Lanka</i>	11	0,06%	Asia	Southern Asia
<i>Sudan</i>	2	0,01%	Africa	Northern Africa
<i>Suriname</i>	1	0,01%	Americas	South America
<i>Sweden</i>	564	3,04%	Europe	Northern Europe
<i>Switzerland</i>	224	1,21%	Europe	Western Europe

<i>Syria</i>	158	0,85%	Asia	Western Asia
<i>Taiwan</i>	19	0,10%	Asia	Eastern Asia
<i>Tajikistan</i>	0	0,00%	Asia	Central Asia
<i>Thailand</i>	57	0,31%	Asia	South-Eastern Asia
<i>The Bahamas</i>	0	0,00%	Americas	Caribbean
<i>Togo</i>	0	0,00%	Africa	Western Africa
<i>Tonga</i>	0	0,00%	Oceania	Polynesia
<i>Trinidad and Tobago</i>	0	0,00%	Americas	Caribbean
<i>Tunisia</i>	7	0,04%	Africa	Northern Africa
<i>Turkey</i>	40	0,22%	Asia	Western Asia
<i>Turkmenistan</i>	0	0,00%	Asia	Central Asia
<i>Tuvalu</i>	0	0,00%	Oceania	Polynesia
<i>Uganda</i>	22	0,12%	Africa	Eastern Africa
<i>Ukraine</i>	8	0,04%	Europe	Eastern Europe
<i>United Arab Emirates</i>	20	0,11%	Asia	Western Asia
<i>United Kingdom</i>	540	2,91%	Europe	Northern Europe
<i>United Republic of Tanzania</i>	34	0,18%	Africa	Eastern Africa
<i>United States of America</i>	2218	11,96%	Americas	Northern America
<i>Uruguay</i>	0	0,00%	Americas	South America
<i>Uzbekistan</i>	0	0,00%	Asia	Central Asia
<i>Vanuatu</i>	0	0,00%	Oceania	Melanesia
<i>Vatican</i>	0	0,00%	Europe	Southern Europe
<i>Venezuela</i>	154	0,83%	Americas	South America
<i>Vietnam</i>	25	0,13%	Asia	South-Eastern Asia
<i>Yemen</i>	28	0,15%	Asia	Western Asia
<i>Zambia</i>	10	0,05%	Africa	Eastern Africa
<i>Zimbabwe</i>	115	0,62%	Africa	Eastern Africa

Abstract

Klika, D. (2022). *Projects related to the SDGs and the response generated on social networks: analysis of the World Economic Forum profile on Twitter over the selected time period* [Master's Thesis, University of West Bohemia].

Key words: sustainability, sustainable development, Sustainable Development Goals, social media analysis, Twitter, World Economic Forum, human geography

This thesis focuses on the analysis of the Twitter profile of the World Economic Forum over the period of 5 years, from 2016 to 2020. The main focus of this analysis is the content posted by the World Economic Forum and its ties to sustainability and Sustainable Development Goals based on multiple approaches. This thesis utilizes freely available programs and custom developed software to acquire and work with its dataset. The data were analyzed and results presented in three categories – geographical content, topical content, and the most popular content. Results presented in this thesis are: a list of countries with the number of mentions in tweets, lists of most popular hashtags and users mentioned by the World Economic Forum, and lists of the most popular tweets based on various interactions (likes, retweets, replies, and quoted retweets). Authors address the possible shortcomings of some parts of the analysis, present their conclusions and propose their ideas for further work utilizing similar input data or methodology.

Abstrakt

Klika, D. (2022). *Projects related to the SDGs and the response generated on social networks: analysis of the World Economic Forum profile on Twitter over the selected time period* [Diplomová práce, Západočeská univerzita v Plzni].

Klíčová slova: udržitelnost, udržitelný rozvoj, Cíle udržitelného rozvoje, analýza sociálních sítí, Twitter, Světové ekonomické fórum, humánní geografie

Tato práce se zaměřuje na analýzu Twitterového profilu Světového ekonomického fóra za období pěti let, od roku 2016 do roku 2020. Hlavním zaměřením této práce je analýza obsahu příspěvků Světového ekonomického fóra a jejich napojení na problematiku udržitelnosti a Cíle udržitelného rozvoje na základě několika přístupů. Tato práce využívá volně dostupné programy společně s vlastním softwarem vyvinutým pro tuto práci k získání datasetu a k práci s ním. Tato data byla analyzována a výsledky prezentovány ve třech kategoriích – geografické zaměření, tematické zaměření a nejpoblárnější obsah. Výsledky prezentované v této práci jsou: seznam zemí a počtu zmínek v tweetech, seznamy nejpoblárnějších hashtagů a uživatelů zmíněných Světovým ekonomickým fórem a seznamy nejpoblárnějších tweetů na základě možných interakcí (like, retweet, odpověď a citovaný retweet). Autoři této práce zohledňují možné nedostatky některých aspektů analýzy, prezentují své závěry a předkládají své myšlenky pro další práci s využitím stejného datasetu nebo metodologie.